

POINTWISE DUAL REPRESENTATION OF DYNAMIC CONVEX EXPECTATIONS

DANIEL BARTL*

ABSTRACT. We fully characterize discrete-time dynamic convex expectations (\mathcal{E}_t) with domain and range the upper semianalytic functions – in particular we work without a reference measure and do not assume essential suprema to exist. It is shown that \mathcal{E}_t is pointwise continuous from below and continuous from above on the continuous functions if and only if a dual representation of \mathcal{E}_t in terms of conditional expectations minus the convex conjugate of \mathcal{E}_t holds true, where the conjugate is lower semianalytic with pointwise weakly compact level sets. Moreover, we provide a dual characterization of the dynamic property, i.e. we show that $\mathcal{E}_t = \mathcal{E}_t \circ \mathcal{E}_{t+1}$ if and only if the convex conjugate of \mathcal{E}_t has an additive form. We also consider dynamic convex expectations defined on the set of discrete-time stochastic processes.

1. INTRODUCTION

A conditional convex expectation is a monotone and convex functional $\mathcal{E}_t: \mathcal{L}_T \rightarrow \mathcal{L}_t$ which is constant on \mathcal{L}_t , where $(\mathcal{L}_t)_{t=0,\dots,T}$ is an increasing family of function spaces. In case of a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ with $\mathcal{F}_0 = \{\emptyset, \Omega\}$, the choice $\mathcal{L}_t := L^\infty(\mathcal{F}_t, P)$ of equivalence classes of bounded and \mathcal{F}_t -measurable random variables yields the classical discrete-time setup and includes in particular the case of static convex expectations $\mathcal{E}_0: L^\infty(\mathcal{F}_T, P) \rightarrow \mathbb{R}$ and up to the sign also convex risk measures. The latter were introduced in [5, 37], and it follows from the Krein-Šmulian theorem that \mathcal{E}_0 admits a dual representation

$$\mathcal{E}_0(X) = \sup_{Q \ll P} (E_Q[X] - \alpha_0(Q))$$

if and only if \mathcal{E}_0 is continuous from below, that is, $\mathcal{E}_0(X_n)$ increases to $\mathcal{E}_0(X)$ whenever $X_n \in L^\infty(\mathcal{F}_T, P)$ is a sequence increasing P -almost surely to $X \in L^\infty(\mathcal{F}_T, P)$, see [28]. In case of conditional convex expectations the characterization reads almost the same: $\mathcal{E}_t: L^\infty(\mathcal{F}_T, P) \rightarrow L^\infty(\mathcal{F}_t, P)$ has the representation

$$\mathcal{E}_t(X) = \operatorname{ess\,sup}_{Q \ll P \text{ and } Q=P \text{ on } \mathcal{F}_t} (E_Q[X|\mathcal{F}_t] - \alpha_t(Q))$$

if and only if $\mathcal{E}_t(X_n)$ increases P -almost surely to $\mathcal{E}_t(X)$ whenever $X_n \in L^\infty(\mathcal{F}_T, P)$ increases P -almost surely to $X \in L^\infty(\mathcal{F}_T, P)$. An important application of this representation is the dual characterization of time-consistency (also called the dynamic property): The family (\mathcal{E}_t) is time-consistent, i.e. one can compute \mathcal{E}_t recursively by

Date: December 30, 2016.

2010 Mathematics Subject Classification. 45N30, 93E20, 46N10, 91B30.

Key words and phrases. nonlinear expectations, dual representation, time-consistency, dynamic programming, robust risk measures for processes, mathematical finance under uncertainty.

*Department of Mathematics, University of Konstanz, daniel.bartl@uni-konstanz.de. The author would like to thank Michael Kupper for fruitful discussions and helpful comments.

$\mathcal{E}_t = \mathcal{E}_t \circ \mathcal{E}_{t+1}$, if and only if α_t has an additive form, see e.g. [2, 3, 22, 33, 38, 49]. In the special case that every \mathcal{E}_t is sublinear, this implies that (\mathcal{E}_t) is time-consistent if and only if \mathcal{E}_t can be expressed as the supremum of conditional expectations over a set of probabilities which is stable under pasting, see e.g. [29]. Similar results can be obtained when \mathcal{L}_t is defined as the set (of equivalence classes) of P -almost surely bounded adapted stochastic processes, see e.g. [2, 19, 20, 21, 22, 39].

Recently, a lot of attention was drawn to convex monotone functionals $\mathcal{E}_0: \mathcal{L}_T \rightarrow \mathbb{R}$, where \mathcal{L}_T is no longer the quotient space $L^\infty(\mathcal{F}_T, P)$ but rather the space of measurable functions from a topological space Ω into the reals. This seems to be mostly due to the relation to robust methods in mathematical finance, where no reference measure P is assumed a priori, see e.g. [1, 17, 32, 34, 40] and [24] for an overview. The absence of P entails several problems, in particular one can not apply classical results from functional analysis such as the Krein-Šmulian theorem in the dual pairing $\langle \mathcal{L}_T, \mathfrak{P} \rangle$, where \mathfrak{P} denotes the set of countably additive probabilities on Ω . As a consequence, no necessary and sufficient sequential continuity conditions for \mathcal{E}_0 to have a dual representation are known, and so far there is essentially only one way to overcome this problem: First one restricts to (semi-) continuous functions in \mathcal{L}_T , where continuity from above is sufficient in order to obtain duality. This condition turns out to be equivalent to weak compactness (in the weak topology induced by the continuous bounded functions C_b) of the level sets $\{\alpha_0 \leq c\}$, see e.g. [1, 11] for a proof via the classical Fenchel-Moreau theorem when \mathcal{E}_0 is given as a superhedging problem and [23, 24] for a general approach based on the Daniell-Stone theorem. The extension to measurable (upper semianalytic) functions requires merely continuity from below on the whole space \mathcal{L}_T and builds on the deep and powerful theory of Choquet [25, 26] on the regularity of (functional) capacities, see e.g. [8, 12, 42]. For the conditional case (i.e. $t > 0$ when $\mathcal{E}_t(X)$ is not a number but a function) the absence of a reference measure reveals further (serious) difficulties: Not only the appropriate continuity of \mathcal{E}_t for a dual representation is unclear, but even the dual objects $E_Q[\cdot | \mathcal{F}_t]$ are defined in an almost sure sense a priori. Moreover, since no essential suprema exist in general, heavy issues regarding measurability occur.

For the rest of the introduction, we assume that $\Omega = \Omega_1^T$ is the Cartesian product of a Polish space Ω_1 and denote by \mathcal{L}_t the space of all bounded upper semianalytic functions from Ω into the reals which only depend on the first t components. By the disintegration theorem, every probability Q on Ω can be written as $Q = \mu \otimes K$ for a probability μ on Ω_1^t and a kernel $K: \Omega_1^t \rightarrow \mathfrak{P}(\Omega_1^{T-t})$ so that $E_Q[X | \mathcal{F}_t](\omega) = E_{K(\omega)}[X(\omega, \cdot)]$ for μ -almost all $\omega \in \Omega_1^t$. The seminal papers [15] and [46] build on this observation. Therein the authors start with sets $(\mathcal{P}_t(\omega))_{t,\omega}$ of probabilities $\mathcal{P}_t(\omega) \subset \mathfrak{P}(\Omega_1)$ and define one-step conditional sublinear expectations $\hat{\mathcal{E}}_t$ by

$$(1) \quad \hat{\mathcal{E}}_t(X)(\omega) := \sup_{Q \in \mathcal{P}_t(\omega)} E_Q[X(\omega, \cdot)] \quad \text{for every } \omega \in \Omega_1^t \text{ and } X \in \mathcal{L}_{t+1}.$$

The profound theory by Luzin and Souslin on analytic sets is then used to show that, under certain measurability assumptions of the family $(\mathcal{P}_t(\omega))_{t,\omega}$, the functional $\hat{\mathcal{E}}_t$ is in fact a mapping from \mathcal{L}_{t+1} to \mathcal{L}_t , i.e. $\mathcal{E}_t(X)$ is upper semianalytic whenever X is. Further, using measurable selection arguments, they show that the constructed one-step expectations $\hat{\mathcal{E}}_t: \mathcal{L}_{t+1} \rightarrow \mathcal{L}_t$ give rise to a dynamic sublinear expectation $\mathcal{E}_t := \hat{\mathcal{E}}_t \circ \dots \circ \hat{\mathcal{E}}_{T-1}$ mapping \mathcal{L}_T to \mathcal{L}_t which can be represented over the set

of probabilities emerging from pasting probabilities in $\mathcal{P}_t, \dots, \mathcal{P}_{T-1}$. Thus, they provide a way of constructing sublinear expectations through a given family of probability measures and show that if the representing set is stable under pasting, then the family (\mathcal{E}_t) is time-consistent.

In this article we study the reverse direction, that is, the starting point here is not a family of probabilities $(\mathcal{P}_t(\omega))_{t,\omega}$ but an arbitrary conditional convex expectation $\mathcal{E}_t: \mathcal{L}_T \rightarrow \mathcal{L}_t$. Our first main result is that, similar as in the classical setting where \mathcal{L}_t equals $L^\infty(\mathcal{F}_t, P)$, the pointwise version of the continuity used in the static case (continuous from below on \mathcal{L}_T and continuous from above on $\mathcal{L}_T \cap C_b$) is equivalent to the validity of a dual representation

$$(2) \quad \mathcal{E}_t(X)(\omega) = \sup_{Q \in \mathfrak{P}(\Omega_1^{T-t})} (E_Q[X(\omega, \cdot)] - \alpha_t(Q, \omega))$$

where the function α_t is jointly lower semianalytic and has pointwise weakly compact level sets $\{\alpha_t(\cdot, \omega) \leq c\}$, see Theorem 2.4. In particular, this implies that under the continuity assumption, every sublinear expectation \mathcal{E}_t has a representation as in (1). Moreover, the family of measures $(\mathcal{P}_t(\omega))_{t,\omega}$ automatically satisfies the assumption considered in [15, 46]. The second main result is the dual characterization of time-consistency: Given that \mathcal{E}_t satisfies the pointwise continuity (and thus has a representation as in (2)), we show that

$$\mathcal{E}_t = \mathcal{E}_t \circ \mathcal{E}_{t+1} \text{ if and only if } \alpha_t(Q, \omega) = \beta_t(Q_t, \omega) + E_{Q_t(d\bar{\omega})}[\alpha_{t+1}(Q_{t+1,T}(\bar{\omega}), (\omega, \bar{\omega}))],$$

where β_t is the convex conjugate of \mathcal{E}_t restricted to \mathcal{L}_{t+1} , and $Q = Q_t \otimes Q_{t+1,T}$ is the disintegration of Q into a measures Q_t on Ω_1 and a kernel $Q_{t+1,T}: \Omega_1 \rightarrow \mathfrak{P}(\Omega_1^{T-t-1})$, see Theorem 2.13. In the second part, we consider convex expectations defined on the set of discrete-time stochastic processes. A one-to-one relation between the same pointwise continuity and a dual representation similar to (2) is given, see Proposition 3.5. Furthermore, dynamic convex expectations for stochastic processes are constructed through given one-step expectations and it is shown that α_t has an additive form, see Theorem 3.4. As a side result, we obtain the disintegration of measures on the optional σ -field as in [2] in the dynamic case. As for the techniques involved, we draw on the theory of analytic sets as well as analytic and Borel selection arguments and the current state of development in representation theory for convex increasing functionals.

Similar questions are studied in [27], however, in a highly different setting: A set of probabilities is fixed and essential suprema (with respect to this set) are assumed to exist, see [44] for a characterization of this property. An additive structure of α_t is also studied in [43], though in a different setting and with a different aim.

The rest of this article is organized as follows: Section 2 starts with defining the present setting and contains all results for convex expectations defined on the space of random variables. In Section 3 we consider convex expectations for (discrete-time) stochastic processes. At the end of both sections robust versions of classical examples illustrate the results. Appendix A contains basic facts about analytic sets and Appendix B the representation of convex increasing functionals.

2. CONVEX EXPECTATIONS FOR RANDOM VARIABLES

We first fix our notation. For a Polish space V , denote by $usa_b(V)$ the set of all upper semianalytic bounded functions from V to \mathbb{R} , by $usc_b(V)$ and $C_b(V)$ the

subsets of upper semicontinuous and continuous functions, respectively. A short summary of analytic (and universally measurable) sets and functions is given in Appendix A. For a function $X: V \rightarrow \mathbb{R}$, write $\|X\|_\infty := \sup_{v \in V} |X(v)|$ for the maximum norm. The set of σ -additive probability measures on the Borel σ -field $\mathcal{B}(V)$ of V is denoted by $\mathfrak{P}(V)$ and endowed with the weak topology $\sigma(\mathfrak{P}(V), C_b(V))$, i.e. the coarsest topology making the mappings $P \mapsto E_P[X]$ continuous for every $X \in C_b(V)$. Then $\mathfrak{P}(V)$ becomes a Polish space itself. Given another Polish space W , the term kernel refers to universally measurable mappings $K: V \rightarrow \mathfrak{P}(W)$. Given such a kernel K and a measure $\mu \in \mathfrak{P}(V)$, the formula $\mu \otimes K(A) := \int_V \int_W 1_A(v, w) K(v)(dw) \mu(dv)$ defines a probability measure $\mu \otimes K$ on $V \times W$. We will always consider the product topology on product spaces and when functions are in consideration, convergence and (in-) equalities are to be understood in a pointwise sense, unless stated otherwise.

2.1. Main results. Fix some $T \in \mathbb{N}$ and a Polish space Ω_1 . Define $\Omega_t := \Omega_1^t$ and

$$\mathcal{L}_t := \text{usab}(\Omega_t) = \{X: \Omega_t \rightarrow \mathbb{R} : X \text{ is upper semianalytic and bounded}\}$$

for $t = 0, 1, \dots, T$ with the convention that Ω_0 is a singleton. Notice that $\mathcal{L}_0 = \mathbb{R}$ and $\mathfrak{P}(\Omega_0) = \{1\}$. For simplicity we will write $\mathcal{L} := \mathcal{L}_T$ and $\Omega := \Omega_T$ and often consider \mathcal{L}_t as a subset of \mathcal{L} , and Ω_t as a subset of Ω .

Definition 2.1. A mapping $\mathcal{E}_t: \mathcal{L} \rightarrow \mathcal{L}_t$ is called *convex expectation at time t* , if for all $X, Y \in \mathcal{L}$,

- $\mathcal{E}_t(X) \leq \mathcal{E}_t(Y)$ whenever $X \leq Y$,
- $\mathcal{E}_t(m) = m$ for all $m \in \mathcal{L}_t$
- $\mathcal{E}_t(\lambda X + (1 - \lambda)Y) \leq \lambda \mathcal{E}_t(X) + (1 - \lambda) \mathcal{E}_t(Y)$ for all $\lambda \in [0, 1]$.

We say that \mathcal{E}_t is a *sublinear expectation at time t* if in addition

- $\mathcal{E}_t(\lambda X) = \lambda \mathcal{E}_t(X)$ for all $\lambda \in [0, +\infty)$.

Let us give a brief comment on the choice of \mathcal{L}_t . Fix some measure $P \in \mathfrak{P}(\Omega)$ and let \mathcal{F}_t denote the σ -field on Ω which is generated by Ω_t . Then $L^\infty(\Omega, \mathcal{F}_t, P)$ coincides with the quotient space \mathcal{L}_t / \sim_P and our definition of convex expectations coincides with the usual one. The same holds true if we were to define \mathcal{L}_t as an arbitrary space in between the Borel and universally measurable functions. However, in the present setup the choice of \mathcal{L}_t turns out to be fundamental and $\mathcal{L}_t = \text{usab}(\Omega_t)$ the only space where a one-to-one dual characterization of \mathcal{E}_t is possible, see Theorem 2.4 and Remark 2.6.

Definition 2.2. Consider the following continuity assumptions on $\mathcal{E}_t: \mathcal{L} \rightarrow \mathcal{L}_t$:

- (A) $\mathcal{E}_t(X_n) \uparrow \mathcal{E}_t(X)$ for every sequence $X_n \in \mathcal{L}$ such that $X_n \uparrow X \in \mathcal{L}$.
- (B) $\mathcal{E}_t(X_n) \downarrow \mathcal{E}_t(X)$ for every sequence $X_n \in C_b(\Omega)$ with $X_n \downarrow X \in \text{usc}_b(\Omega)$.

We will give a comment on both conditions in Remark 2.6. The following useful dual characterization of condition (B) is worth mentioning.

Lemma 2.3. \mathcal{E}_t satisfies (B) if and only if \mathcal{E}_t satisfies both

- (B') $\mathcal{E}_t(X_n) \downarrow 0$ for every sequence $X_n \in C_b(\Omega)$ with $X_n \downarrow 0$,
- (B'') $\sup_{X \in C_b(\Omega)} (E_Q[X] - \mathcal{E}_t(X)) = \sup_{X \in \text{usc}_b(\Omega)} (E_Q[X] - \mathcal{E}_t(X))$ for $Q \in \mathfrak{P}(\Omega)$.

Proof. Since Ω is a Polish space, every upper semicontinuous function can be written as the decreasing limit of a sequence of continuous functions. Therefore it is

clear that (B) implies (B') and (B''). The other direction will be shown within the proof of Theorem 2.4. \square

For every $t = 0, \dots, T-1$, $\omega \in \Omega_t$, and $Q \in \mathfrak{P}(\Omega_{T-t})$ define the so-called penalty function

$$(3) \quad \alpha_t(Q, \omega) := \sup_{X \in \mathcal{C}_b(\Omega)} (E_Q[X(\omega, \cdot)] - \mathcal{E}_t(X)(\omega)).$$

Then the following theorem holds.

Theorem 2.4. *Let \mathcal{E}_t be a convex expectation at time t which satisfies (A) and (B). Then for every $\omega \in \Omega_t$, and $X \in \mathcal{L}$, it holds*

$$(4) \quad \mathcal{E}_t(X)(\omega) = \sup_{Q \in \mathfrak{P}(\Omega_{T-t})} (E_Q[X(\omega, \cdot)] - \alpha_t(Q, \omega))$$

and α_t given by (3) is lower semianalytic, convex in Q^1 , $\inf_Q \alpha_t(Q, \omega) = 0$, and $\{\alpha_t(\cdot, \omega) \leq c\}$ is compact for every $\omega \in \Omega_t$ and $c \in \mathbb{R}$.

Conversely, if $\alpha_t: \mathfrak{P}(\Omega_{T-t}) \times \Omega_t \rightarrow [0, +\infty]$ is a given function which is lower semianalytic and satisfies $\inf_Q \alpha_t(Q, \omega) = 0$ for every ω , then \mathcal{E}_t defined by (4) is a convex expectation at time t satisfying (A). If in addition α_t is convex in Q and $\{\alpha_t(\cdot, \omega) \leq c\}$ is compact for every c and ω , then \mathcal{E}_t satisfies also (B).

Corollary 2.5. *Let \mathcal{E}_t be a sublinear expectation at time t which satisfies (A) and (B). Then for every $\omega \in \Omega_t$, and $X \in \mathcal{L}$, it holds*

$$(5) \quad \mathcal{E}_t(X)(\omega) = \sup_{Q \in \mathcal{Q}_t(\omega)} E_Q[X(\omega, \cdot)]$$

for a set-valued mapping $\mathcal{Q}_t: \Omega_t \rightsquigarrow \mathfrak{P}(\Omega_{T-t})$ with non-empty convex compact values and analytic graph.

Conversely, if $\mathcal{Q}_t: \Omega_t \rightsquigarrow \mathfrak{P}(\Omega_{T-t})$ is a set-valued mapping with non-empty values and analytic graph, then \mathcal{E}_t given by (5) defines a sublinear expectation at time t satisfying (A). If additionally \mathcal{Q}_t has compact convex values, then \mathcal{E}_t satisfies (B).

Proof. If \mathcal{E}_t is sublinear, a scaling argument shows that α_t defined by (3) only takes the values 0 and $+\infty$, and one can define $\mathcal{Q}_t(\omega) := \{\alpha_t(\cdot, \omega) = 0\}$. Conversely, if \mathcal{Q}_t is given, just set $\alpha_t(Q, \omega) := +\infty 1_{\mathcal{Q}_t(\omega)^c}(Q)$. \square

Remark 2.6. (i) Notice that whenever a dual representation of \mathcal{E}_t as in (4) is possible, then condition (A) necessarily holds true. This follows by interchanging two suprema and the monotone convergence theorem. To the best of our knowledge there are, even for $t = 0$, no conditions on sequential continuity of \mathcal{E}_t which improve (B) in order to obtain a dual representation. By Theorem 2.4 this condition corresponds to compactness of the level sets of α_t , which in applications is often satisfied (see e.g. [11, 24]).

(ii) Let \mathcal{E}_t be a convex expectation at time t which satisfies (A) and (B). Even if $\mathcal{E}_t(X)$ is continuous whenever X is continuous and α_t in (4) is independent of ω , $\mathcal{E}_t(X)$ needs not to be Borel for X Borel. Moreover, it is not possible to modify \mathcal{E}_t on a \mathcal{E}_{t-1} -polar set² in order to obtain a Borel measurable version. In general one can not define \mathcal{E}_t on a linear space which contains all bounded Borel functions.

¹That is, for every ω the mapping $Q \mapsto \alpha_t(Q, \omega)$ is convex.

²That is, a set which has measure zero for every Q which satisfy $\alpha_{t-1}(Q, \omega) < +\infty$.

Proof. The following example builds on the fact that projections of Borel sets need not to be Borel and is similar to [46, Chapter 5.2]. Let $T = 2$, $\Omega_1 = [0, 1]$, and define

$$\mathcal{E}_0(X) := \sup_{Q \in \mathcal{Q}_0} E_Q[X] \quad \text{and} \quad \mathcal{E}_1(X)(\omega) := \sup_{Q \in \mathcal{Q}_1(\omega)} E_Q[X(\omega, \cdot)]$$

where $\mathcal{Q}_0 := \mathcal{Q}_1(\omega) := \mathfrak{P}([0, 1])$ for $\omega \in \Omega_1$. It is straightforward to show that \mathcal{E}_t is a sublinear expectation which fulfills (A) and (B) and that $\mathcal{E}_1(X)$ is continuous whenever X is continuous. Let $B \subset \Omega$ be a Borel set which projection on the first component (denoted by A) is not Borel. Then $\mathcal{E}_1(1_B) = 1_A$ which is not Borel. Further, since all Dirac measures appear in the set \mathcal{Q}_0 , only the empty set is polar and we can not modify 1_A at all. Using the same arguments, one can copy [46, Chapter 5.3] in order to prove the second statement in (ii). \square

Remark 2.7. *In the static case, i.e. for $t = 0$, it is possible to show that (B') is equivalent to the better known "tightness" condition: There exists a sequence of compact sets $K_n \subset \Omega$ such that $\mathcal{E}_0(r1_{K_n^c}) \downarrow 0$ for every $r > 0$. In general, i.e. for $t > 0$, this no longer holds true.*

Proof. Here we only provide a counterexample, the equivalence of (B') and the tightness condition in the static case is shown within the proof of Theorem 2.4. Let $T = 2$ and define

$$\mathcal{E}_1(X)(\omega) := X(\omega, \omega) \quad \text{for } \omega \in \Omega_1 := \mathbb{R}^{\mathbb{N}} \text{ and } X \in \mathcal{L},$$

where $\mathbb{R}^{\mathbb{N}}$ is endowed with the product topology. It is straightforward to show that \mathcal{E}_1 is a sublinear expectation which satisfies (A) and (B). Assume that there exists a (increasing) sequence of compact sets $K_n \subset \Omega$ such that $\mathcal{E}_1(1_{K_n^c}) \downarrow 0$ pointwise, where we may assume without loss of generality that $K_n = C_n \times C_n$ for $C_n \subset \Omega_1$ compact. Since $\mathcal{E}_1(1_{K_n^c})(\omega) = 1_{C_n^c}(\omega)$ it follows that $\Omega_1 = \bigcup \{C_n : n\}$ and thus, by the Baire category theorem, there exists some n such that C_n has non-empty interior. However, it is straightforward to show that every compact subset of Ω_1 has empty interior, and thus such a sequence K_n cannot exist. \square

Proof of Theorem 2.4. (a) Let \mathcal{E}_t be a convex expectation at time t which satisfies (A) and (B) and fix some t and $\omega \in \Omega_t$. From the part of Lemma 2.3 which was already proven and Lemma B.2, it follows that the functional $\mathcal{E}_t(\cdot)(\omega) : \mathcal{L} \rightarrow \mathbb{R}$ satisfies all assumptions of Theorem B.1 with $M = C_b(\Omega)$. Thus

$$\mathcal{E}_t(X)(\omega) = \sup_{\bar{Q} \in \mathfrak{M}(M)} (E_{\bar{Q}}[X] - \mathcal{E}_t^*(\bar{Q}, \omega)) \quad \text{for all } X \in \mathcal{L}$$

where $\mathcal{E}_t^*(\bar{Q}, \omega) := \sup_{X \in C_b(\Omega)} (E_{\bar{Q}}[X] - \mathcal{E}_t(X)(\omega))$ and $\mathfrak{M}(M)$ denotes the set of all positive measures on the σ -field generated by M , i.e. the Borel σ -field. First, we claim that every \bar{Q} with $\mathcal{E}_t^*(\bar{Q}, \omega) < +\infty$ satisfies $\bar{Q} = \delta_\omega \otimes Q$ for some $Q \in \mathfrak{P}(\Omega_{T-t})$, where δ_ω denotes the Dirac measure assigning probability 1 to the point ω . By assumption it holds $\mathcal{E}_t(m)(\omega) = m(\omega)$ for every $m \in C_b(\Omega_t)$. Therefore $E_{\bar{Q}}[m] = m(\omega)$, otherwise a scaling argument yields $\mathcal{E}_t^*(\bar{Q}, \omega) = +\infty$. Taking all constant functions m , it follows that \bar{Q} has to be a probability measure. Now, by the disintegration theorem, we may write $\bar{Q} = \mu \otimes K$ for a measure $\mu \in \mathfrak{P}(\Omega_t)$ and a kernel $K : \Omega_t \rightarrow \mathfrak{P}(\Omega_{T-t})$. Since $E_\mu[m] = E_{\bar{Q}}[m] = m(\omega)$ for all $m \in C_b(\Omega_t)$, it holds $\mu = \delta_\omega$ and thus $Q := K(\omega)$ satisfies $\bar{Q} = \delta_\omega \otimes Q$. Define

$$\alpha_t(Q, \omega) := \mathcal{E}_t^*(\delta_\omega \otimes Q, \omega)$$

for $Q \in \mathfrak{P}(\Omega_{T-t})$ and notice that $\inf_Q \alpha_t(Q, \omega) = 0$ since $\mathcal{E}_t(0)(\omega) = 0$.

We are left to show that α_t is lower semianalytic and convex in Q . Since Ω is a Polish space, there exists a metric d' on Ω which induces the original topology under which the space $uc_b(\Omega, d')$ becomes separable (see e.g. [51, Lemma 3.1.4]). Here $uc_b(\Omega, d')$ denotes the set of all bounded functions from Ω to \mathbb{R} which are uniformly continuous with respect to d' , and this space is endowed with the maximum norm. Let D' be a countable dense subset and define

$$D := \{(-n) \vee X \wedge n : X \in D', n \in \mathbb{N}\} \subset C_b(\Omega)$$

so that D is still countable. Now fix $\omega \in \Omega_t$, $Q \in \mathfrak{P}(\Omega_{T-t})$, $X \in C_b(\Omega)$, $\varepsilon > 0$, and let m be a natural number which is larger than $\|X\|_\infty$. Since the set $\Lambda_{2m}(\omega) := \{\alpha_t(\cdot, \omega) \leq 2m\}$ is compact by Theorem B.1, Prokhorov's theorem yields the existence of a compact subset C of Ω_{T-t} such that

$$Q(C^c) \leq \frac{\varepsilon}{m} \quad \text{and} \quad \sup_{P \in \Delta_{2m}(\omega)} P(C^c) \leq \frac{\varepsilon}{m}.$$

Then $K := \{\omega\} \times C$ is a compact subset of Ω and since the metric d' induces the original topology on Ω , the set $K \subset (\Omega, d')$ is still compact and $X : (\Omega, d') \rightarrow \mathbb{R}$ still continuous. Thus $X1_K \in uc_b(K, d')$ and by a generalization of Tietze's extension theorem (see [41, Theorem 3]) there exists a uniformly continuous function $Y \in uc_b(\Omega, d')$ such that $Y = X$ on K . Since $D' \subset uc_b(\Omega, d')$ is dense, there exists $Z' \in D'$ such that $\|Z' - Y\|_\infty \leq \varepsilon$. Let $Z := (-m) \vee Z' \wedge m \in D$ so that

$$\begin{aligned} E_P[Z(\omega, \cdot)] &\leq E_P[X(\omega, \cdot)1_C(\cdot)] + \varepsilon + E_P[Z(\omega, \cdot)1_{C^c}(\cdot)] \\ &\leq E_P[X(\omega, \cdot)1_C(\cdot)] + E_P[X(\omega, \cdot)1_{C^c}(\cdot)] + 3\varepsilon = E_P[X(\omega, \cdot)] + 3\varepsilon \end{aligned}$$

for every $P \in \Lambda_{2m}(\omega)$. Changing the roles of X and Z and replacing P by Q yields $E_Q[X(\omega, \cdot)] \leq E_Q[Z(\omega, \cdot)] + 3\varepsilon$. By monotonicity of \mathcal{E}_t it holds $-\|Y\|_\infty \leq \mathcal{E}_t(Y)(\omega) \leq \|Y\|_\infty$ for every $Y \in \mathcal{L}$, thus

$$(6) \quad \mathcal{E}_t(Y)(\omega) = \sup_{P \in \Lambda_{2m}(\omega)} (E_P[Y(\omega, \cdot)] - \alpha_t(P, \omega)) \quad \text{for } Y \in \mathcal{L} \text{ with } \|Y\|_\infty \leq m.$$

This implies $\mathcal{E}_t(Z)(\omega) \leq \mathcal{E}_t(X)(\omega) + 3\varepsilon$ and therefore

$$\begin{aligned} E_Q[X(\omega, \cdot)] - \mathcal{E}_t(X)(\omega) &\leq E_Q[Z(\omega, \cdot)] - \mathcal{E}_t(Z)(\omega) + 6\varepsilon \\ &\leq \sup_{Z \in D} (E_Q[Z(\omega, \cdot)] - \mathcal{E}_t(Z)(\omega)) + 6\varepsilon. \end{aligned}$$

Since D is a subset of $C_b(\Omega)$ and $X \in C_b(\Omega)$, $\varepsilon > 0$ were arbitrary, it follows that

$$\alpha_t(Q, \omega) = \sup_{Z \in D} (E_Q[Z(\omega, \cdot)] - \mathcal{E}_t(Z)(\omega)).$$

Finally, for every $Z \in D$, the function $(Q, \omega) \mapsto E_Q[Z(\omega, \cdot)] - \mathcal{E}_t(Z)(\omega)$ is lower semianalytic (see [14, Proposition 7.29]) and convex in Q . As the countable supremum, α_t inherits both properties.

(b) Let $\alpha_t : \mathfrak{P}(\Omega_{T-t}) \times \Omega_t \rightarrow [0, +\infty]$ be a given lower semianalytic function which satisfies $\inf_Q \alpha_t(Q, \omega) = 0$ for every ω , and define

$$\mathcal{E}_t(X)(\omega) := \sup_{Q \in \mathfrak{P}(\Omega_{T-t})} (E_Q[X(\omega, \cdot)] - \alpha_t(Q, \omega))$$

for $\omega \in \Omega_t$ and $X \in \mathcal{L}$. Since $X \in \mathcal{L}$ is upper semianalytic, the mapping

$$\mathfrak{P}(\Omega_{T-t}) \times \Omega_t \rightarrow [-\infty, +\infty), \quad (Q, \omega) \mapsto E_Q[X(\omega, \cdot)] - \alpha_t(Q, \omega)$$

is upper semianalytic (see [14, Proposition 7.48]) and it follows from [14, Proposition 7.47] that $\omega \mapsto \mathcal{E}_t(X)(\omega)$ is upper semianalytic. Further, since $\inf_Q \alpha_t(Q, \omega) = 0$, we have that

$$-\|X\|_\infty \leq \mathcal{E}_t(X)(\omega) \leq \|X\|_\infty.$$

Thus $\mathcal{E}_t(0) = 0$ and $\mathcal{E}_t(X)$ is bounded, therefore $\mathcal{E}_t(X) \in \mathcal{L}_t$. The other properties needed for \mathcal{E}_t to be a convex expectation at time t are immediate. Further, condition (A) follows by interchanging two suprema and the monotone convergence theorem.

Assume in addition that α_t is convex in Q and $\Lambda_c(\omega) := \{\alpha_t(\cdot, \omega) \leq c\}$ is compact for every real number c and $\omega \in \Omega_t$. Fix some $\omega \in \Omega_t$, $X_\infty \in usc_b(\Omega)$, and let $X_n \in C_b(\Omega)$ be a sequence which decreases pointwise to X_∞ . Then it follows as in (6) that

$$(7) \quad \mathcal{E}_t(X_n)(\omega) = \max_{Q \in \Lambda_c(\omega)} (E_Q[X_n(\omega, \cdot)] - \alpha_t(Q, \omega)) \quad \text{for } n \in \mathbb{N} \cup \{\infty\},$$

where $c := 2(\|X_1\|_\infty \vee \|X_\infty\|_\infty)$. Since $\Lambda_c(\omega)$ is compact and convex, X_n is a decreasing sequence, and

$$\mathfrak{P}(\Omega_{T-t}) \ni Q \mapsto E_Q[X_n(\omega, \cdot)] - \alpha_t(Q, \omega)$$

is convex and upper semicontinuous for every $n \in \mathbb{N}$, it follows from (7), a minimax theorem (see [36, Theorem 2]), and the monotone convergence theorem that

$$\inf_{n \in \mathbb{N}} \mathcal{E}_t(X_n)(\omega) = \max_{Q \in \Lambda_c(\omega)} \inf_{n \in \mathbb{N}} (E_Q[X_n(\omega, \cdot)] - \alpha_t(Q, \omega)) = \mathcal{E}_t(X_\infty)(\omega).$$

Thus (B) and also the missing part of Lemma 2.3 are proven. \square

Definition 2.8. A family $(\mathcal{E}_t)_{0 \leq t \leq T}$ is called a *dynamic (convex) expectation*, if \mathcal{E}_t is a convex expectation at time t for $t = 0, \dots, T$ and

- $\mathcal{E}_t(X) = \mathcal{E}_t(\mathcal{E}_{t+1}(X))$ for every $X \in \mathcal{L}$ and $t = 0, \dots, T-1$.

We call a family $(\varphi_t)_{0 \leq t \leq T-1}$ of mappings $\varphi_t: \mathcal{L}_{t+1} \rightarrow \mathcal{L}_t$ a *generator*, if every φ_t is a convex expectation at time t according to Definition 2.1 with $T = t+1$. The analogous terminology is adopted when we say that (φ_t) satisfies (A) and (B).

Any generator (φ_t) induces a dynamic expectation (\mathcal{E}_t) through the formula

$$(8) \quad \mathcal{E}_T(X) := X \quad \text{and} \quad \mathcal{E}_t := \varphi_t \circ \dots \circ \varphi_{T-1} \quad \text{for every } t = 0, \dots, T-1.$$

Conversely, given a family of convex expectation (\mathcal{E}_t) , we may define φ_t as the restriction of \mathcal{E}_t to \mathcal{L}_{t+1} . If (\mathcal{E}_t) is dynamic, then $\mathcal{E}_t = \varphi_t \circ \dots \circ \varphi_{T-1}$. Moreover, whenever (φ_t) satisfies (A) and (B), it follows from Theorem 2.4 that

$$(9) \quad \begin{cases} \varphi_t(X)(\omega) = \sup_{Q \in \mathfrak{P}(\Omega_1)} (E_Q[X(\omega, \cdot)] - \beta_t(Q, \omega)) \text{ for every } X \in \mathcal{L}_{t+1} \\ \text{where } \beta_t: \mathfrak{P}(\Omega_1) \times \Omega_t \rightarrow [0, +\infty] \text{ is lower semianalytic, } \inf_Q \beta_t(Q, \omega) = 0 \end{cases}$$

for $t = 0, \dots, T-1$ and $\omega \in \Omega_t$. Conversely, if the representation (9) holds true, it follows again from Theorem 2.4 that (φ_t) defines a generator.

Proposition 2.9. Fix a generator (φ_t) which satisfies (A) and (B) or, more generally, assume that (9) holds, and let (\mathcal{E}_t) be the associated dynamic expectation defined by (8). Then it holds

$$(10) \quad \mathcal{E}_t(X)(\omega) := \sup_{Q \in \mathfrak{P}(\Omega_{T-t})} (E_Q[X(\omega, \cdot)] - \alpha_t(Q, \omega))$$

for $t = 0, \dots, T-1$, $\omega \in \Omega_t$, and $X \in \mathcal{L}$, for the function α_t defined as

$$(11) \quad \alpha_t(Q, \omega) := \sum_{s=t}^{T-1} E_Q[\beta_s(Q_s(\cdot), (\omega, \cdot))]$$

where $Q_s: \Omega_{s-t} \rightarrow \mathfrak{P}(\Omega_1)$ are kernels³ such that $Q = Q_t \otimes \dots \otimes Q_{T-1}$.

Corollary 2.10. Fix a sublinear generator (φ_t) which satisfies (A) and (B) or, more generally, let $\mathcal{Q}_t: \Omega_t \rightsquigarrow \mathfrak{P}(\Omega_1)$ be set-valued mappings with non-empty values and analytic graph such that $\varphi_t(X)(\omega) = \sup_{Q \in \mathcal{Q}_t(\omega)} E_Q[X(\omega, \cdot)]$ for $X \in \mathcal{L}_{t+1}$ and $t = 0, \dots, T-1$. The associated dynamic sublinear expectation (\mathcal{E}_t) defined by (8) satisfies

$$\mathcal{E}_t(X)(\omega) = \sup_{Q \in \mathcal{Q}_t^T(\omega)} E_Q[X(\omega, \cdot)]$$

for $t = 0, \dots, T-1$, $\omega \in \Omega_t$, and $X \in \mathcal{L}$ for

$$\mathcal{Q}_t^T(\omega) := \{Q_t \otimes \dots \otimes Q_{T-1} : Q_s(\cdot) \in \mathcal{Q}_s(\omega, \cdot)\}$$

where $Q_s: \Omega_{s-t} \rightarrow \mathfrak{P}(\Omega_1)$ are kernels such that $Q_s(\bar{\omega}) \in \mathcal{Q}_s(\omega, \bar{\omega})$ for $\bar{\omega} \in \Omega_{s-t}$.

Remark 2.11. One can show that α_t defined by (11) is lower semianalytic. This follows from the fact that the disintegration of a measure can be achieved by countably many operations, thus preserving measurability, see e.g. [14, Proposition 7.27].

Remark 2.12. Fix a dynamic expectation (\mathcal{E}_t) with associated generator (φ_t) . If every \mathcal{E}_t satisfies (A) and (B), then so does (φ_t) . The converse does not hold.

Proof. The first statement follows from the definition. We give a counterexample in which (φ_t) satisfies (A) and (B), but (\mathcal{E}_t) does not. Let $T = 2$, $\Omega_1 = \mathbb{R}$ and define

$$\varphi_0(X) := \sup_{\omega \in [0,1]} X(\omega) \quad \text{and} \quad \varphi_1(X)(\omega) := X\left(\omega, \frac{1}{\omega}\right)$$

with the convention that $1/0 = 0$. Since $\omega \mapsto (\omega, 1/\omega)$ is Borel, it follows from [14, Lemma 7.30] that $\omega \mapsto \varphi_1(X)(\omega)$ is upper semianalytic. Therefore (φ_t) is a (sublinear) generator and one can verify that (A) and (B) are satisfied. Define the sequence of functions

$$X_n(\omega_1, \omega_2) := (\omega_2 - n + 1)1_{[n-1, n]}(\omega_2) + 1_{(n, \infty)}(\omega_2) \in C_b(\Omega)$$

which satisfies $X_n \downarrow 0$ but

$$\mathcal{E}_0(X_n) := \varphi_0 \circ \varphi_1(X_n) = \sup_{\omega \in [0,1]} X_n\left(\omega, \frac{1}{\omega}\right) \geq X_n\left(\frac{1}{n}, n\right) = 1.$$

Hence \mathcal{E}_0 does not satisfy (B). \square

Proof of Proposition 2.9. Fix $X \in \mathcal{L}$. We show by a backward induction over t that

(i) given any $\varepsilon > 0$, there exists a kernel $Q: \Omega_t \rightarrow \mathfrak{P}(\Omega_{T-t})$ such that

$$\mathcal{E}_t(X)(\omega) \leq E_{Q(\omega)}[X(\omega, \cdot)] - \alpha_t(Q(\omega), \omega) + \varepsilon$$

for every $\omega \in \Omega_t$,

(ii) (10) holds for every $\omega \in \Omega_t$.

³With slight abuse of notation, we identify the kernel $Q_t: \Omega_{t-t} = \Omega_0 = \{\omega_0\} \rightarrow \mathfrak{P}(\Omega_1)$ with the measure $Q_t(\omega_0) \in \mathfrak{P}(\Omega_1)$. In particular $Q := Q_t \otimes \dots \otimes Q_{T-1}$ is a measures and not a kernel.

For $t = T - 1$ we have $\beta_{T-1} = \alpha_{T-1}$ so that there is nothing to prove for claim (ii). Moreover, since

$$\mathfrak{P}(\Omega_1) \times \Omega_{T-1} \rightarrow [-\infty, +\infty), \quad (Q, \omega) \mapsto E_Q[X(\omega, \cdot)] - \beta_{T-1}(Q, \omega)$$

is upper semianalytic by [14, Proposition 7.48] and X is bounded from above, claim (i) follows from the Jankov-von Neumann theorem (see [14, Proposition 7.50]). Now assume that both claims hold true for $t + 1$ and let $R: \Omega_{t+1} \rightarrow \mathfrak{P}(\Omega_{T-t-1})$ be a kernel such that

$$\mathcal{E}_{t+1}(X)(\omega) \leq E_{R(\omega)}[X(\omega, \cdot)] - \alpha_{t+1}(R(\omega), \omega) + \varepsilon/2$$

for every $\omega \in \Omega_{t+1}$, where ε is as usual arbitrary but chosen in advance. Further, since $\mathcal{E}_t = \varphi_t \circ \mathcal{E}_{t+1}$ by definition and $\mathcal{E}_{t+1}(X) \in \mathcal{L}_{t+1}$, the same arguments used for $t = T - 1$ yield the existence of a kernel $P: \Omega_t \rightarrow \mathfrak{P}(\Omega_1)$ such that

$$\varphi_t(\mathcal{E}_{t+1}(X))(\omega) \leq E_{P(\omega)}[\mathcal{E}_{t+1}(X)(\omega, \cdot)] - \beta_t(P(\omega), \omega) + \varepsilon/2$$

for every $\omega \in \Omega_t$. Define $Q(\omega) := P(\omega) \otimes R(\omega, \cdot)$ for $\omega \in \Omega_t$. It follows from [14, Lemma 7.29] that $Q(\omega)$ is well-defined and e.g. from [14, Lemma 7.28] and a twofold application of [14, Proposition 7.46] that $Q: \Omega_t \rightarrow \mathfrak{P}(\Omega_{T-t})$ is a kernel. But then we have that

$$\begin{aligned} \mathcal{E}_t(X)(\omega) &= \varphi_t(\mathcal{E}_{t+1}(X))(\omega) \leq E_{P(\omega)}[\mathcal{E}_{t+1}(X)(\omega, \cdot)] - \beta_t(P(\omega), \omega) + \varepsilon/2 \\ &\leq E_P[E_{R(\omega, \cdot)}[X(\omega, \cdot)] - \alpha_{t+1}(R(\omega, \cdot), (\omega, \cdot))] - \beta_t(P(\omega), \omega) + \varepsilon \\ &= E_{Q(\omega)}[X(\omega, \cdot)] - \alpha_t(Q(\omega), \omega) + \varepsilon \end{aligned}$$

for $\omega \in \Omega_t$ which shows claim (i) and in particular that the left hand side in (10) is smaller than the right hand side. Thus, in order to prove claim (ii), we are left to show the reverse inequality. Fix $\omega \in \Omega_t$ and a measure $Q \in \mathfrak{P}(\Omega_{T-t})$, which we shall write as $Q = P \otimes R$ for a measure $P \in \mathfrak{P}(\Omega_1)$ and a kernel $R: \Omega_1 \rightarrow \mathfrak{P}(\Omega_{T-t-1})$. Then

$$\begin{aligned} \mathcal{E}_t(X)(\omega) &= \varphi_t(\mathcal{E}_{t+1}(X))(\omega) \geq E_P[E_{R(\cdot)}[X(\omega, \cdot)] - \alpha_{t+1}(R(\cdot), (\omega, \cdot))] - \beta_t(P, \omega) \\ &= E_Q[X(\omega, \cdot)] - \alpha_t(Q, \omega) \end{aligned}$$

and as Q was arbitrary, this implies that the left hand side in (10) is larger than the right hand side, which completes the proof. \square

In Proposition 2.9 we have shown that a dynamic expectations can be represented by a penalty function which equals the sum over the one-step penalties. The following theorem shows that, given additional regularity, this penalty is also the so-called minimal penalty function (resp. the convex conjugate). Define

$$(12) \quad \alpha_{t,s}(Q, \omega) := \sup_{X \in C_b(\Omega_s)} (E_Q[X(\omega, \cdot)] - \mathcal{E}_t(X)(\omega))$$

for every $t < s$, $Q \in \mathfrak{P}(\Omega_{s-t})$, and $\omega \in \Omega_t$.

Theorem 2.13. *For every t , let \mathcal{E}_t be a convex expectation at time t which satisfies (A) and (B). Further assume that $\mathcal{E}_t(X)$ is Borel for every $X \in C_b(\Omega)$. Then (\mathcal{E}_t) is a dynamic expectation if and only if*

$$(13) \quad \alpha_{t,s}(Q, \omega) = \alpha_{t,r}(P, \omega) + E_Q[\alpha_{r,s}(R(\cdot), (\omega, \cdot))]$$

for every $t < r < s$ and $\omega \in \Omega_t$, where each α is given by (12) and $Q = P \otimes R$ for a measure $P \in \mathfrak{P}(\Omega_{r-t})$ and a kernel $R: \Omega_{r-t} \rightarrow \mathfrak{P}(\Omega_{s-r})$.

For the sublinear case, define

$$(14) \quad \mathcal{Q}_{t,s}(\omega) := \{Q \in \mathfrak{P}(\Omega_{s-t}) : E_Q[X(\omega, \cdot)] \leq \mathcal{E}_t(X)(\omega) \text{ for all } X \in C_b(\Omega_s)\} \\ = \{Q \in \mathfrak{P}(\Omega_{s-t}) : \alpha_{t,s}(Q, \omega) = 0\}$$

for every $t < s$ and $\omega \in \Omega_t$.

Corollary 2.14. *For every t , let \mathcal{E}_t be a sublinear expectation at time t which satisfies (A) and (B). Further assume that $\mathcal{E}_t(X)$ is Borel for every $X \in C_b(\Omega)$. Then (\mathcal{E}_t) is a dynamic sublinear expectation if and only if*

$$(15) \quad \mathcal{Q}_{t,s}(\omega) = \mathcal{Q}_{t,r}(\omega) \otimes \mathcal{Q}_{r,s}(\omega, \cdot)$$

for every $t < s < r$ and $\omega \in \Omega_t$, where each \mathcal{Q} is given by (14) and the right hand side in (15) is to be understood as the set of all measures $P \otimes R$ where $P \in \mathcal{Q}_{t,r}(\omega)$ and the kernel $R: \Omega_{r-t} \rightarrow \mathfrak{P}(\Omega_{s-r})$ satisfies $R(\bar{\omega}) \in \mathcal{Q}_{r,s}(\omega, \bar{\omega})$ for P -almost all $\bar{\omega} \in \Omega_{r-t}$.

Proof of Theorem 2.13. If (13) holds, then it follows from Proposition 2.9 that (\mathcal{E}_t) is a dynamic expectation. Conversely, if (\mathcal{E}_t) is a dynamic expectation such that every \mathcal{E}_t satisfies (A) and (B), it again follows from Proposition 2.9 that

$$\mathcal{E}_t(X)(\omega) = \sup_{Q=P \otimes R \in \mathfrak{P}(\Omega_{s-t})} (E_Q[X(\omega, \cdot)] - (\alpha_{t,r}(P, \omega) + E_P[\alpha_{r,s}(R(\cdot), (\omega, \cdot))]))$$

for $X \in \mathcal{L}_s$ and in particular that

$$\alpha_{t,s}(Q, \omega) = \sup_{X \in C_b(\Omega_s)} (E_Q[X(\omega, \cdot)] - \mathcal{E}_t(X)(\omega)) \\ \leq \alpha_{t,r}(P, \omega) + E_P[\alpha_{r,s}(R(\cdot), (\omega, \cdot))].$$

We are left to show the other inequality. Fix some t , $\omega \in \Omega_t$, and $Q \in \mathfrak{P}(\Omega_{s-t})$, and observe that

$$(16) \quad \alpha_{t,s}(Q, \omega) = \sup\{E_Q[X] : X \in C_b(\Omega_{s-t}) \text{ such that } \mathcal{E}_t(X \circ \pi_{t,s})(\omega) \leq 0\}$$

$$(17) \quad = \sup\{E_Q[X] : X \in \text{usab}(\Omega_{s-t}) \text{ such that } \mathcal{E}_t(X \circ \pi_{t,s})(\omega) \leq 0\}$$

where $\pi_{t,s}: \Omega_s \rightarrow \Omega_{s-t}$ denotes the canonical projection on the last $s-t$ components. Indeed let $Y \in C_b(\Omega_s)$ as in the definition of $\alpha_{t,s}$ and set $X := Y(\omega, \cdot) - \mathcal{E}_t(Y)(\omega)$. Then, using the dual representation of \mathcal{E}_t , we obtain that

$$E_Q[X] = E_Q[Y(\omega, \cdot)] - \mathcal{E}_t(Y)(\omega) \quad \text{and} \quad \mathcal{E}_t(X \circ \pi_{t,s})(\omega) = 0$$

which implies (16). By definition (16) is smaller than (17) and since on the other hand $E_Q[X] - \alpha_{t,s}(Q, \omega) \leq 0$ for every $X \in \text{usab}(\Omega_{s-t})$ such that $\mathcal{E}_t(X \circ \pi_{t,s})(\omega) \leq 0$, it follows that (17) actually equals (16). From now on, fix $t < r < s$, let $P \in \mathfrak{P}(\Omega_{r-t})$ and $R: \Omega_{r-t} \rightarrow \mathfrak{P}(\Omega_{s-r})$ be Borel such that $Q = P \otimes R$. In the sequel we will always

write $\bar{\omega}$ for elements in Ω_{r-t} and $\bar{\bar{\omega}}$ for elements in Ω_{s-r} .

We may argue as in the proof of Theorem 2.4 and choose a metric d' under which the space of bounded and uniformly continuous functions $uc_b(\Omega_{s-r}, d')$ becomes separable. Let D' be a countable dense subset of $uc_b(\Omega_{s-r}, d')$ and define $D = \{(-m) \vee d \wedge m : d \in D', m \in \mathbb{N}\}$. Further let $\{d_n : n\}$ be an enumeration of D and $\{q_n : n\}$ an enumeration of the positive rational numbers, and define

$$D_n := \{0\} \cup \{d_k - q_l : 1 \leq k, l \leq n\} \subset C_b(\Omega_{s-r})$$

for every natural number n . We claim that

$$(18) \quad \alpha_{r,s}(R(\bar{\omega}), (\omega, \bar{\omega})) = \sup_n \alpha_{r,s}^n(R(\bar{\omega}), (\omega, \bar{\omega}))$$

for every $\bar{\omega} \in \Omega_{r-s}$, where, for every natural number n ,

$$\alpha_{r,s}^n(R(\bar{\omega}), (\omega, \bar{\omega})) := \max\{E_{R(\bar{\omega})}[X] : X \in D_n \text{ such that } \mathcal{E}_r(X \circ \pi_{r,s})(\omega, \bar{\omega}) \leq 0\}.$$

By (16) (applied to “ $t = r$ ”) we have that the left hand side in (18) is larger than the right hand side. To show the reverse inequality fix some $\bar{\omega} \in \Omega_{r-t}$ and $X \in C_b(\Omega_{r-s})$ such that $\mathcal{E}_r(X \circ \pi_{r,s})(\omega, \bar{\omega}) \leq 0$. Then, by the same arguments as in the proof of Theorem 2.4, there exists a sequence $X_n \in D$ such that

$$E_{R(\bar{\omega})}[X_n] \rightarrow E_{R(\bar{\omega})}[X] \quad \text{and} \quad \mathcal{E}_r(X_n \circ \pi_{r,s})(\omega, \bar{\omega}) \rightarrow \mathcal{E}_r(X \circ \pi_{r,s})(\omega, \bar{\omega}) \leq 0.$$

Thus, for any rational number $\varepsilon > 0$ and n_0 large enough, it holds $E_{R(\bar{\omega})}[X_{n_0}] \geq E_{R(\bar{\omega})}[X] - \varepsilon$ and $\mathcal{E}_r(X_{n_0} \circ \pi_{r,s})(\omega, \bar{\omega}) \leq \varepsilon$. Define $Y := X_{n_0} - \varepsilon$ so that $E_{R(\bar{\omega})}[Y] \geq E_{R(\bar{\omega})}[X] - 2\varepsilon$ and $\mathcal{E}_r(Y \circ \pi_{r,s})(\omega, \bar{\omega}) \leq 0$. Since $X_{n_0} \in D$ and $\varepsilon > 0$ is rational, we have that $Y \in D_n$ for some large n . As $\varepsilon > 0$ was arbitrary, (18) is established. Further $\alpha_{r,s}^n \geq 0$ since $0 \in D_n$, and $\alpha_{r,s}^n \leq \alpha_{r,s}^{n+1}$ since $D_n \subset D_{n+1}$. Therefore the monotone convergence theorem applies and yields

$$\alpha_{t,r}(P, \omega) + E_P[\alpha_{r,s}(R(\cdot), (\omega, \cdot))] = \sup_n \sup_Y (E_P[Y] + E_P[\alpha_{r,s}^n(R(\cdot), (\omega, \cdot))])$$

where the supremum is taken over all $Y \in C_b(\Omega_{r-t})$ satisfying $\mathcal{E}_t(Y \circ \pi_{t,r})(\omega) \leq 0$, see (16). For the rest of the proof,

$$\text{fix } n \in \mathbb{N} \quad \text{and} \quad Y \in C_b(\Omega_{r-t}) \text{ such that } \mathcal{E}_t(Y \circ \pi_{t,r})(\omega) \leq 0.$$

In what follows we construct a function $Z \in \text{usab}(\Omega_{s-t})$ which satisfies

$$E_P[Y] + E_P[\alpha_{r,s}^n(R(\cdot), (\omega, \cdot))] = E_Q[Z] \quad \text{and} \quad \mathcal{E}_t(Z \circ \pi_{t,s})(\omega) \leq 0.$$

Then, as n and Y were arbitrary, by (17), we can conclude that

$$\alpha_{t,s}(Q, \omega) \geq \alpha_{t,r}(P, \omega) + E_P[\alpha_{r,s}(R(\cdot), (\omega, \cdot))]$$

and the proof is finished. Define the set-valued mapping

$$\Psi(\bar{\omega}) := \{X \in D_n : E_{R(\bar{\omega})}[X] = \alpha_{r,s}^n(R(\bar{\omega}), (\omega, \bar{\omega})) \text{ and } \mathcal{E}_r(X \circ \pi_{r,s})(\omega, \bar{\omega}) \leq 0\}.$$

We endow the set D_n with the discrete topology⁴ which, since D_n is a finite set, makes D_n a Polish space and claim that Ψ is weakly measurable, that is,

$$\Psi^l(O) := \{\bar{\omega} : \Psi(\bar{\omega}) \cap O \neq \emptyset\} \quad \text{is Borel}$$

for every open subset O of D_n . First notice that, since every $X \in D_n$ is continuous and bounded, the mapping $\bar{\omega} \mapsto \mathcal{E}_r(X \circ \pi_{r,s})(\omega, \bar{\omega})$ is Borel by assumption. Moreover $Q \mapsto E_Q[X]$ is continuous which implies that $\bar{\omega} \mapsto E_{R(\bar{\omega})}[X]$ is Borel. Therefore

$$\alpha_{r,s}^n(R(\bar{\omega}), (\omega, \bar{\omega})) = \max_{X \in D_n} (E_{R(\bar{\omega})}[X] - \infty 1_{(0, \infty)}(\mathcal{E}_r(X \circ \pi_{r,s})(\omega, \bar{\omega})))$$

is Borel as a function of $\bar{\omega}$ which implies that

$$\Psi^l(O) = \bigcup_{X \in O} \{\bar{\omega} : E_{R(\bar{\omega})}[X] = \alpha_{r,s}^n(R(\bar{\omega}), (\omega, \bar{\omega})) \text{ and } \mathcal{E}_r(X \circ \pi_{r,s})(\omega, \bar{\omega}) \leq 0\}$$

is Borel, as the finite union of Borel sets. Since $0 \in D_n$, it follows that Ψ has non-empty values, and since the topology on D_n is the discrete one, the values of

⁴That is, all subsets of D_n are open and closed.

Ψ are closed as well. Therefore we may apply the Kuratowski–Ryll–Nardzewski theorem (see e.g. [4, Theorem 18.13]) and obtain a Borel measurable mapping

$$X^\Psi: \Omega_{r-t} \rightarrow D_n \quad \text{such that} \quad X^\Psi(\bar{\omega}) \in \Psi(\bar{\omega}).$$

Now define the function

$$Z(\bar{\omega}, \bar{\bar{\omega}}) := Y(\bar{\omega}) + X^\Psi(\bar{\omega})(\bar{\bar{\omega}}) \quad \text{for } (\bar{\omega}, \bar{\bar{\omega}}) \in \Omega_{r-t} \times \Omega_{s-r} = \Omega_{s-t}.$$

For every $\bar{\omega}$, the mapping $Z(\bar{\omega}, \cdot)$ is continuous by definition. Moreover, since D_n is armed with the discrete topology, the mapping $Z(\cdot, \bar{\bar{\omega}})$ is Borel for every $\bar{\bar{\omega}}$. A basic fact on Carathéodory-functions (see e.g. [4, Lemma 4.51]) yields that

$$Z \quad \text{is jointly Borel and bounded,}$$

where the boundedness follows since D_n is a finite set consisting of bounded functions. Moreover

$$\mathcal{E}_r(Z \circ \pi_{t,s})(\omega, \bar{\omega}) = Y(\bar{\omega}) + \mathcal{E}_r(X^\Psi(\bar{\omega}) \circ \pi_{r,s})(\omega, \bar{\omega}) \leq Y(\bar{\omega})$$

so that

$$\mathcal{E}_t(Z \circ \pi_{t,s})(\omega) = \mathcal{E}_t(\mathcal{E}_r(Z \circ \pi_{t,s}))(\omega) \leq \mathcal{E}_t(Y \circ \pi_{t,r})(\omega) \leq 0$$

by the dynamic property and monotonicity of \mathcal{E}_t . Further it holds

$$E_Q[Z] = E_P[Y] + E_{P(d\bar{\omega})}[E_{R(\bar{\omega})}[X^\Psi(\bar{\omega})]] = E_P[Y] + E_{P(d\bar{\omega})}[\alpha_{r,s}^n(R(\bar{\omega}), (\omega, \bar{\omega}))]$$

which completes the proof. \square

2.2. Examples. The aim of this subsection is to show how one can apply the previous results to construct robust dynamic risk measures through well-known classical ones. For every t , let $\mathcal{P}_t(\omega) \subset \mathfrak{P}(\Omega_1)$ be a convex non-empty set of possible probability scenarios for the next time period (given the path $\omega \in \Omega_t$) such that the set-valued mapping $\omega \mapsto \mathcal{P}_t(\omega)$ has analytic graph. This is the setting proposed in [15] and adopted e.g. in [9, 18, 45]. Further define the set of possible probability scenarios for the whole future (given the path $\omega \in \Omega_t$)

$$(19) \quad \mathcal{P}_t^T(\omega) = \{P_t \otimes \cdots \otimes P_{T-1} : P_s(\cdot) \in \mathcal{P}_s(\omega, \cdot) \text{ for } s = t, \dots, T-1\}.$$

For any two probability measures Q and P we denote by dQ/dP the Radon–Nikodym derivate of the absolutely continuous part of Q with respect to P and write $Q \ll P$ if Q is absolutely continuous with respect to P .

2.2.1. Average value at risk. See also [3, 6, 22] for the classical (non-robust) version of this example. Let $\lambda_t: \Omega_t \rightarrow (0, 1)$ be Borel and define

$$(20) \quad \varphi_t(X)(\omega) = \frac{1}{\lambda_t(\omega)} \inf_{s \in \mathbb{R}} \left(\sup_{P \in \mathcal{P}_t(\omega)} E_P[(X(\omega, \cdot) + s)^+] - s\lambda_t(\omega) \right)$$

for $X \in \mathcal{L}_{t+1}$.

Example 2.15. The family (φ_t) forms a generator and the associated dynamic expectation $\mathcal{E}_t := \varphi_t \circ \cdots \circ \varphi_{T-1}$ has the representation

$$\mathcal{E}_t(X)(\omega) = \sup_{Q \in \mathcal{Q}_t^T(\omega)} E_Q[X(\omega, \cdot)]$$

for $\omega \in \Omega_t$ and $X \in \mathcal{L}$, where

$$\mathcal{Q}_s(\omega) = \{Q \in \mathfrak{P}(\Omega_1) : Q \ll P \text{ and } dQ/dP \leq 1/\lambda_s(\omega) \text{ for some } P \in \mathcal{P}_s(\omega)\}$$

for every s and $\omega \in \Omega_s$, and $\mathcal{Q}_t^T(\omega)$ is defined as in (19).

Proof. Fix t and ω . The term $E_P[(X(\omega, \cdot) + s)^+] - s\lambda_t(\omega)$ tends to $+\infty$ whenever $|s|$ does. Thus, in (20), it suffices to minimize over s in some compact subset of \mathbb{R} regardless of whether the order of optimization is $\inf_s \sup_P$ or $\sup_P \inf_s$. Therefore it follows from a minimax theorem (see [36, Theorem 2]), the dual representation for the classical average value at risk (see [37, Lemma 4.51 and Theorem 4.52]) and by interchanging two suprema that $\varphi_t(X)(\omega) = \sup_{Q \in \mathcal{Q}_t(\omega)} E_Q[X(\omega, \cdot)]$. Further the function f defined by

$$f(\omega, Q, P) := E_P[1_{[0, 1/\lambda_t(\omega)]}(dQ/dP)] E_P[dQ/dP].$$

is Borel by [31, Theorem V.58] and [14, Proposition 7.29]. Since the graph of \mathcal{P}_t is analytic, it follows that

$$\text{graph } \mathcal{Q}_t = \pi\{(\omega, Q, P) : P \in \mathcal{P}_t(\omega) \text{ and } f(\omega, Q, P) \geq 1\}$$

is an analytic set, where π denotes the projection on the first two components. Hence the claim follows from Corollary 2.10. \square

2.2.2. Entropic risk measure. See also [3, 22]. Recall that the relative entropy $H(Q, P)$ of Q with respect to P is defined as

$$H(Q, P) := \begin{cases} E_P[\frac{dQ}{dP} \log \frac{dQ}{dP}], & \text{if } Q \ll P, \\ +\infty, & \text{otherwise.} \end{cases}$$

In analogy we define the robust entropy $H(Q, \mathcal{P}) := \inf_{P \in \mathcal{P}} H(Q, P)$ for a set of probability measures \mathcal{P} . For every t , let $\lambda_t : \Omega_t \rightarrow (0, 1]$ be Borel such that $1/\lambda_t$ is bounded, and define

$$\varphi_t(X)(\omega) := \frac{1}{\lambda_t(\omega)} \sup_{P \in \mathcal{P}_t(\omega)} \log E_P[\exp(\lambda_t(\omega)X(\omega, \cdot))]$$

for $X \in \mathcal{L}_{t+1}$.

Example 2.16. The family (φ_t) forms a generator and the associated dynamic expectation $\mathcal{E}_t := \varphi_t \circ \dots \circ \varphi_{T-1}$ has the representation

$$\mathcal{E}_t(X)(\omega) = \sup_{Q \in \mathfrak{P}(\Omega_{T-t})} \left(E_Q[X(\omega, \cdot)] - \sum_{s=t}^{T-1} E_Q \left[\frac{1}{\lambda_s(\omega, \cdot)} H(Q_s(\cdot), \mathcal{P}_s(\omega, \cdot)) \right] \right)$$

for $X \in \mathcal{L}$ and $\omega \in \Omega_t$. If in addition $\lambda_t = \lambda$ for some $\lambda \in (0, 1]$ and every t , then

$$(21) \quad \mathcal{E}_t(X)(\omega) = \frac{1}{\lambda} \sup_{P \in \mathcal{P}_t^T(\omega)} \log E_P[\exp(\lambda X(\omega, \cdot))]$$

$$(22) \quad = \sup_{Q \in \mathfrak{P}(\Omega_{T-t})} \left(E_Q[X(\omega, \cdot)] - \frac{1}{\lambda} H(Q, \mathcal{P}_t^T(\omega)) \right).$$

Proof. By interchanging two suprema, it follows from the classical representation of the expected exponential (see e.g. [37, Lemma 3.29]) that

$$\varphi_t(X)(\omega) = \sup_{Q \in \mathfrak{P}(\Omega_1)} \left(E_Q[X(\omega, \cdot)] - \frac{1}{\lambda_t(\omega)} H(Q, \mathcal{P}_t(\omega)) \right).$$

Moreover $H(P, \mathcal{P}_t(\omega)) = 0$ for every $P \in \mathcal{P}_t(\omega)$ and from [7, Lemma 3.9 and Lemma 2.5] we have that $(Q, \omega) \mapsto H(Q, \mathcal{P}_t(\omega))$ is lower semianalytic and convex in Q , so that the first claim follows from Proposition 2.9. If additionally $\lambda_t = \lambda$, it follows from selection arguments in line with the ones from Proposition 2.9 that (21) holds true and (22) follows again by [7, Lemma 3.9]. \square

3. CONVEX EXPECTATIONS FOR STOCHASTIC PROCESSES

3.1. Main results. Fix some $T \in \mathbb{N}$ and a Polish space Ω_1 . For $t = 0, \dots, T$ define $\Omega_t := \Omega_1^t$ and $\mathcal{L}_t := \text{usab}(\Omega_t)$, with the convention that Ω_0 is a singleton, and

$$\mathcal{R}_t := \{X_t 1_{\{t\}} + \dots + X_T 1_{\{T\}} : X_s \in \mathcal{L}_s \text{ for } s = t, \dots, T\}.$$

Then \mathcal{R}_t corresponds to the set of (adapted) processes starting at time t . Thus we can, and often will,

identify \mathcal{R}_t with $\mathcal{L}_t \times \dots \times \mathcal{L}_T$.

In this spirit, we write $\mathcal{R}_t^{C_b}$ for the set $C_b(\Omega_t) \times \dots \times C_b(\Omega_T)$ and $\mathcal{R}_t^{\text{usc}_b}$ for the set $\text{usc}_b(\Omega_t) \times \dots \times \text{usc}_b(\Omega_T)$.

For two processes X and Y in \mathcal{R}_t , we write $X \leq Y$ if $X_s(\omega) \leq Y_s(\omega)$ for every $\omega \in \Omega_s$ and $s = t, \dots, T$. The multiplication of a process $X \in \mathcal{R}_t$ and a function $m \in \mathcal{L}_t$ is defined pointwise, i.e. $(mX)_s := mX_s$. In contrast to the previous section, we shall use superscripts for sequences of stochastic processes. The following definition of convex expectations for processes can be checked to coincide with the classical one, as soon as a reference probability is fixed, see Section 2.

Definition 3.1. A mapping $\mathcal{E}_t: \mathcal{R}_t \rightarrow \mathcal{L}_t$ is called *convex expectation* (for processes) at time t , if for all $X, Y \in \mathcal{R}_t$

- $\mathcal{E}_t(X) \leq \mathcal{E}_t(Y)$ whenever $X \leq Y$,
- $\mathcal{E}_t(m 1_{[t, T]}) = m$ for all $m \in \mathcal{L}_t$,
- $\mathcal{E}_t(\lambda X + (1 - \lambda)Y) \leq \lambda \mathcal{E}_t(X) + (1 - \lambda) \mathcal{E}_t(Y)$ for all $\lambda \in [0, 1]$.

We call a family of convex expectations $(\mathcal{E}_t)_{0 \leq t \leq T}$ a *dynamic (convex) expectation*, if \mathcal{E}_t is a convex expectation at time t for every t and

- $\mathcal{E}_t(X) = \mathcal{E}_t(X_t 1_{\{t\}} + \mathcal{E}_{t+1}(X 1_{[t+1, T]}) 1_{[t+1, T]})$ for every $t = 0, \dots, T - 1$.

Similar as in the previous section, a family $(\varphi_t)_{0 \leq t \leq T-1}$ of mappings $\varphi_t: \mathcal{L}_t \times \mathcal{L}_{t+1} \rightarrow \mathcal{L}_t$ is called *generator*, if every φ_t is a convex expectation at time t according to Definition 3.1 with $T = t + 1$. Any generator (φ_t) gives rise to a convex expectation \mathcal{E}_t by

$$(23) \quad \mathcal{E}_T(X) := X \text{ and } \mathcal{E}_t(X) := \varphi_t(X_t, \mathcal{E}_{t+1}(X 1_{[t+1, T]})) \text{ for } t = 0, \dots, T - 1.$$

Conversely, for a family of convex expectation (\mathcal{E}_t) we may define φ_t as the restriction of \mathcal{E}_t to $\mathcal{L}_t \times \mathcal{L}_{t+1}$. If (\mathcal{E}_t) is dynamic, then \mathcal{E}_t coincides with the expectation defined by (φ_t) through (23).

Definition 3.2. Consider the following continuity assumptions on $\mathcal{E}_t: \mathcal{L} \rightarrow \mathcal{L}_t$:

- (A) $\mathcal{E}_t(X^n) \uparrow \mathcal{E}_t(X)$ for every sequence $X^n \in \mathcal{R}_t$ such that $X^n \uparrow X \in \mathcal{R}_t$.
- (B) $\mathcal{E}_t(X^n) \downarrow \mathcal{E}_t(X)$ for every sequence $X^n \in \mathcal{R}_t^{C_b}$ with $X^n \downarrow X \in \mathcal{R}_t^{\text{usc}_b}$.

The same proof as in Lemma 2.3 yields a dual characterization of (B).

Lemma 3.3. \mathcal{E}_t satisfies (B) if and only if

- (B') $\mathcal{E}_t(X^n) \downarrow 0$ for all sequences $X^n \in \mathcal{R}_t^{C_b}$ with $X^n \downarrow 0$,
- (B'') $\sup_{X \in \mathcal{R}_t^{C_b}} (E_{\bar{Q}}[X] - \mathcal{E}_t(X)) = \sup_{X \in \mathcal{R}_t^{\text{usc}_b}} (E_{\bar{Q}}[X] - \mathcal{E}_t(X))$ for $\bar{Q} \in \mathfrak{P}(\bar{\Omega})$

where $\bar{\Omega} := \{t, \dots, T\} \times \Omega$.

Before we are ready to state the main results of this section, one last definition is needed. For $t = 0, \dots, T$ define

$$\Gamma_t := \left\{ \gamma = (\gamma_t, \dots, \gamma_T) : \begin{array}{l} \gamma_s : \Omega_{s-t} \rightarrow [0, 1] \text{ is universally measurable} \\ \text{and } \sum_{s=t}^T \gamma_s(\omega_{t+1}, \dots, \omega_s) = 1 \text{ for every } \omega \in \Omega_{T-t} \end{array} \right\}$$

and for every $\gamma \in \Gamma_t$ define

$$d_t^\gamma := 1 \quad \text{and} \quad d_s^\gamma := 1 - \gamma_t - \dots - \gamma_{s-1}$$

for $s = t+1, \dots, T$. Then d^γ is a positive predictable decreasing process starting at 1 and can be seen as a discounting process. The representation (25) which we will prove in Theorem 3.4 can be seen as a decomposition of the uncertainty modeled by a probability Q on the optional σ -field on $\{t, \dots, T\} \times \Omega$ into model uncertainty on Ω (modeled by a probability Q on Ω) and discounting uncertainty (modeled by $\gamma \in \Gamma$). For an extensive discussion see [2], in particular Theorem 3.4 and Appendix B therein. By Proposition 3.5 any generator (φ_t) which satisfies (A) and (B) has the dual representation

$$(24) \quad \begin{cases} \varphi_t(X_t, X_{t+1})(\omega) = \sup_{Q \in \mathfrak{P}(\Omega_1), q \in [0, 1]} (qX_t(\omega) + (1-q)E_Q[X_{t+1}(\omega, \cdot)] - \beta_t(Q, q, \omega)), \\ \beta_t : \mathfrak{P}(\Omega_1) \times [0, 1] \times \Omega_t \rightarrow [0, +\infty] \text{ is lower semianalytic, } \inf_{Q, q} \beta_t(Q, q, \omega) = 0 \end{cases}$$

for every $(X_t, X_{t+1}) \in \mathcal{L}_t \times \mathcal{L}_{t+1}$, $\omega \in \Omega_t$, and $t = 0, \dots, T-1$. Conversely, given that every φ_t has the representation (24), it follows again from Proposition 3.5 that (φ_t) defines a generator.

Theorem 3.4. *Fix a generator (φ_t) which satisfies (A) and (B) or, more generally, assume that (24) holds, and let (\mathcal{E}_t) be the corresponding dynamic expectation defined as in (23). Then \mathcal{E}_t satisfies (A) (but not necessarily (B)) and it holds*

$$(25) \quad \mathcal{E}_t(X)(\omega) = \sup_{Q \in \mathfrak{P}(\Omega_{T-t}), \gamma \in \Gamma_t} \left(\sum_{s=t}^T E_Q[\gamma_s(\cdot)X_s(\omega, \cdot)] - \alpha_t(Q, \gamma, \omega) \right)$$

for $t = 0, \dots, T-1$, $\omega \in \Omega_t$, and $X \in \mathcal{R}_t$, for the function α_t defined by

$$\alpha_t(Q, \gamma, \omega) := \sum_{s=t}^{T-1} E_Q \left[d_s^\gamma(\cdot) \beta_s \left(Q_s(\cdot), \frac{\gamma_s(\cdot)}{d_s^\gamma(\cdot)}, (\omega, \cdot) \right) \right]$$

where $Q_s : \Omega_{s-t} \rightarrow \mathfrak{P}(\Omega_1)$ are kernels such that $Q = Q_t \otimes \dots \otimes Q_{T-1}$.

The proof relies on the static version, which we shall state and prove first. Define $\mathfrak{M}_{\leq 1}$ to be the set of all sub-probabilities, that is, positive measures with total mass smaller than 1 and endow this space again with weak topology $\sigma(\mathfrak{M}_{\leq 1}, C_b)$. This topology makes $\mathfrak{M}_{\leq 1}$ a Polish space and it behaves much like the weak topology on \mathfrak{P} , for example Prokhorov's theorem holds true in the same form: A subset $C \subset \mathfrak{M}_{\leq 1}$ is relatively compact if and only if for every $\varepsilon > 0$ there exists a compact set K such that $\sup_{\mu \in C} \mu(K^c) \leq \varepsilon$, see e.g. [16, Chapter IX]. For $t = 0, \dots, T-1$ define the set

$$\mathcal{A}_t := \{ \mu \in \mathfrak{M}_{\leq 1}(\Omega_0) \times \dots \times \mathfrak{M}_{\leq 1}(\Omega_{T-t}) : \mu_t(\Omega_0) + \dots + \mu_T(\Omega_{T-t}) = 1 \}$$

where $\mu = (\mu_t, \dots, \mu_T)$. Then \mathcal{A}_t , as a closed subspace of the product of Polish spaces, is a Polish space itself. Further define

$$(26) \quad \hat{\alpha}_t(\mu, \omega) := \sup_{X \in \mathcal{R}_t^{C_b}} \left(\sum_{s=t}^T E_{\mu_s}[X_s(\omega, \cdot)] - \mathcal{E}_t(X)(\omega) \right)$$

for $\mu \in \mathcal{A}_t$ and $\omega \in \Omega_t$.

Proposition 3.5. *Let \mathcal{E}_t be a convex expectation at time t which satisfies (A) and (B). Then, for every $X \in \mathcal{R}_t$ and $\omega \in \Omega_t$, it holds*

$$(27) \quad \mathcal{E}_t(X)(\omega) = \sup_{\mu \in \mathcal{A}_t} \left(\sum_{s=t}^T E_{\mu_s}[X_s(\omega, \cdot)] - \hat{\alpha}_t(\mu, \omega) \right),$$

where $\hat{\alpha}_t : \mathcal{A}_t \times \Omega_t \rightarrow [0, +\infty]$ given by (26) is lower semianalytic, convex in μ , satisfies $\inf_{\mu} \hat{\alpha}_t(\mu, \omega) = 0$, and $\{\hat{\alpha}_t(\cdot, \omega) \leq c\}$ is compact for all $\omega \in \Omega_t$ and $c \in \mathbb{R}$.

Conversely, if $\hat{\alpha}_t : \mathcal{A}_t \times \Omega_t \rightarrow [0, +\infty]$ is a given lower semianalytic function which satisfies $\inf_{\mu} \hat{\alpha}_t(\mu, \omega) = 0$ for all ω , then \mathcal{E}_t given by (27) defines a convex expectation at time t which satisfies (A). If in addition α_t is convex in μ and $\{\hat{\alpha}_t(\cdot, \omega) \leq c\}$ is compact for every c , then \mathcal{E}_t satisfies (B).

Proof. We keep the proof short because it is essentially the same as the one given in Theorem 2.4. Define $\bar{\Omega} := \{t, \dots, T\} \times \Omega$ where $\{t, \dots, T\}$ is endowed with the discrete topology, as well as

$$M := \{X \in C_b(\bar{\Omega}) : X_s(\omega) \text{ depends only on } (\omega_1, \dots, \omega_s) \text{ for } s = t, \dots, T\}.$$

We can identify $\mathcal{R}_t^{C_b}$ with M and, in the notation of Theorem B.1, also $\mathcal{R}_t^{usc_b}$ with $M_{\delta, b}$ and \mathcal{R}_t with $S(M)_b$ by Lemma B.2. Now fix some $\omega \in \Omega_t$ so that $\mathcal{E}_t(\cdot)(\omega)$ can be viewed as a mapping from $S(M)_b$ to \mathbb{R} . By assumption (A) and (B) we may apply Theorem B.1 and obtain

$$\mathcal{E}_t(X)(\omega) = \sup_{\bar{Q} \in \mathfrak{M}(M)} (E_{\bar{Q}}[X] - \mathcal{E}_t^*(\bar{Q}, \omega)) \quad \text{for } X \in S(M)_b$$

where $\mathfrak{M}(M)$ denotes the set of all countably additive measures on $(\bar{\Omega}, \sigma(M))$ and $\mathcal{E}_t^*(\bar{Q}, \omega) := \sup_{X \in M} (E_{\bar{Q}}[X] - \mathcal{E}_t(X)(\omega))$. A straightforward computation shows that $\sigma(M)$, the smallest σ -field on $\bar{\Omega}$ which makes all functions $X \in M$ measurable, is given by

$$\sigma(M) = \sigma(\{s\} \times (A \times \Omega_{T-s}) : A \in \mathcal{B}(\Omega_s), s = t, \dots, T).$$

Since $\mathcal{E}_t(x1_{[t, T]})(\omega) = x$ for every $x \in \mathbb{R}$, it follows that $\mathcal{E}_t^*(\bar{Q}, \omega) = +\infty$ whenever \bar{Q} is not a probability measure. Now fix some $\bar{Q} \in \mathfrak{M}(M)$ and define the measures $\mu'_s \in \mathfrak{M}_{\leq 1}(\Omega_s)$ by $\mu'_s := \bar{Q}(\{s\} \times (\cdot \times \Omega_{T-s}))$ for $s = t, \dots, T$ so that

$$(28) \quad E_{\bar{Q}}[X] = \sum_{s=t}^T E_{\mu'_s}[X_s] \quad \text{for every } X \in M.$$

Finally, whenever $\mathcal{E}_t^*(\bar{Q}, \omega) < +\infty$, we have that $\mu'_s = \delta_{\omega} \otimes \mu_s$ for some $\mu \in \mathcal{A}_t$, for if not, it is possible to find $m \in C_b(\Omega_t)$ such that $\sum_{s=t}^T E_{\mu'_s}[m] - m(\omega) > 0$. Since $\mathcal{E}(m1_{[t, T]})(\omega) = m(\omega)$ a scaling argument then implies $\mathcal{E}_t^*(\bar{Q}, \omega) = +\infty$. For $\mu \in \mathcal{A}_t$ define $\bar{Q}^\mu \in \mathfrak{M}(M)$ by (28) where $\mu'_s := \delta_{\omega} \otimes \mu_s$. Then

$$\hat{\alpha}_t(\mu, \omega) = \sup_{X \in M} \left(\sum_{s=t}^T E_{\mu_s}[X_s(\omega, \cdot)] - \mathcal{E}_t(X)(\omega) \right) = \mathcal{E}_t^*(\bar{Q}^\mu, \omega)$$

and therefore (27) holds. The rest of the proof follows as in Theorem 2.4. \square

Proof of Theorem 3.4. If the generator (φ_t) satisfies (A) and (B), then it follows from Proposition 3.5 that (24) holds. Indeed, Proposition 3.5 guarantees that φ_t has the dual representation (27) with $T = t + 1$. Moreover \mathcal{A}_{T-1} can be identified with the set of tuples $(q, Q) \in [0, 1] \times \mathfrak{M}_{\leq 1}(\Omega_1)$ such that $q + Q(\Omega_1) = 1$ and one can define $\beta_t(Q, q, \omega) := \hat{\alpha}_t(q, (1 - q)Q, \omega)$. Conversely, if every φ_t has the representation (24), then one can define $\hat{\alpha}_{T-1}(\mu, \omega) := \beta_t(\mu_{T-1}, \mu_T / (1 - \mu_{T-1}), \omega)$ with $T = t + 1$ and the convention $0/0 := \delta_{\omega_0}$ for some fixed $\omega_0 \in \Omega_{t+1}$. Then φ_t has the representation (27) and Proposition 3.5 yields that (φ_t) defines a generator.

For every t define the set

$$\Gamma'_t := \left\{ \gamma = (\gamma_t, \dots, \gamma_T) : \begin{array}{l} \gamma_s : \Omega_s \rightarrow [0, 1] \text{ is universally measurable} \\ \text{and } \sum_{s=t}^T \gamma_s(\omega_1, \dots, \omega_s) = 1 \text{ for every } \omega \in \Omega \end{array} \right\}.$$

Then $\gamma(\omega, \cdot) \in \Gamma_t$ for all $\gamma \in \Gamma'_t$ and $\omega \in \Omega_t$ by [14, Lemma 7.29]. In what follows we show by a backward induction over t that

- (i) given $\varepsilon > 0$ and $X \in \mathcal{R}_t$, there exists $\gamma \in \Gamma'_t$ and a kernel $Q : \Omega_t \rightarrow \mathfrak{P}(\Omega_{T-t})$ such that

$$\mathcal{E}_t(X)(\omega) \leq \sum_{s=t}^T E_{Q(\omega)}[\gamma_s(\omega, \cdot) X_s(\omega, \cdot)] - \alpha_t(Q(\omega), \gamma(\omega, \cdot), \omega) + \varepsilon$$

for every $\omega \in \Omega_t$,

- (ii) (25) holds for every $\omega \in \Omega_t$ and $X \in \mathcal{R}_t$.

Then, from claim (ii) and the monotone convergence theorem it follows that \mathcal{E}_t satisfies (A) and an example as in Remark 2.12 shows that (B) is not necessarily satisfied.

First notice that Γ_{T-1} can be identified with $[0, 1]$ by $\gamma_{T-1} = q$ and $\gamma_T = 1 - q$. Therefore $\alpha_{T-1}(Q, \gamma, \omega) = \beta_{T-1}(Q, q, \omega)$ and since

$$\begin{aligned} \mathcal{E}_{T-1}(X)(\omega) &= \varphi_{T-1}(X_{T-1}, X_T)(\omega) \\ &= \sup_{(Q, q) \in \mathfrak{P}(\Omega_1) \times [0, 1]} (E_Q[qX_{T-1}(\omega) + (1 - q)X_T(\omega, \cdot)] - \beta_{T-1}(Q, q, \omega)) \end{aligned}$$

claim (ii) follows for $t = T - 1$. Further, since

$$(\omega, Q, q) \mapsto qX_{T-1}(\omega) + (1 - q)E_Q[X_T(\omega, \cdot)] - \beta_{T-1}(Q, q, \omega)$$

is upper semianalytic, it follows that for every $\varepsilon > 0$ there exists a universally measurable mapping $\omega \mapsto (q(\omega), Q(\omega))$ such that

$$\mathcal{E}_{T-1}(X)(\omega) \leq q(\omega)X_{T-1}(\omega) + (1 - q(\omega))E_{Q(\omega)}[X_T(\omega, \cdot)] - \beta_{T-1}(Q(\omega), q(\omega), \omega) + \varepsilon$$

(see [14, Proposition 7.50]). In particular

$$\omega \mapsto q(\omega) \quad \text{and} \quad \omega \mapsto Q(\omega) \quad \text{are universally measurable.}$$

Define $\gamma_{T-1}(\omega) := q(\omega)$ for $\omega \in \Omega_{T-1}$ and $\gamma_T(\omega) := 1 - q(\omega_1, \dots, \omega_{T-1})$ for $\omega \in \Omega$. Then γ_T is universally measurable (see e.g. [14, Proposition 7.44]) so that $\gamma \in \Gamma'_{T-1}$ and claim (i) follows for $t = T - 1$.

Now assume that (i) and (ii) are true for $t + 1$ and let $\varepsilon > 0$ be arbitrary. Since

$$\mathcal{E}_t(X)(\omega) = \sup_{(Q, q) \in \mathfrak{P}(\Omega_1) \times [0, 1]} (qX_t(\omega) + (1 - q)E_Q[\mathcal{E}_{t+1}(X1_{[t+1, T]})(\omega, \cdot)] - \beta_t(Q, q, \omega)),$$

by the same arguments as above, there exist universally measurable mappings $q(\omega)$ and $P(\omega)$ such that

$$\begin{aligned}\mathcal{E}_t(X)(\omega) &\leq q(\omega)X_t(\omega) + (1 - q(\omega))E_{P(\omega)}[\mathcal{E}_{t+1}(X1_{[t+1,T]})(\omega, \cdot)] \\ &\quad - \beta_t(P(\omega), q(\omega), \omega) + \varepsilon/2\end{aligned}$$

for every $\omega \in \Omega_t$. By (i) there exists a kernel $R: \Omega_{t+1} \rightarrow \mathfrak{P}(\Omega_{T-t-1})$ and $\eta \in \Gamma'_{t+1}$ such that

$$\mathcal{E}_{t+1}(X1_{[t+1,T]})(\omega) \leq \sum_{s=t+1}^T E_{R(\omega)}[\eta_s(\omega, \cdot)X_s(\omega, \cdot)] - \alpha_{t+1}(R(\omega), \eta(\omega, \cdot), \omega) + \varepsilon/2$$

for every $\omega \in \Omega_{t+1}$. Define

$$\begin{aligned}\gamma_t(\omega) &:= q(\omega) \quad \text{for } \omega \in \Omega_t, \\ \gamma_s(\omega) &:= (1 - q(\omega_1, \dots, \omega_t))\eta_s(\omega) \quad \text{for } \omega \in \Omega_s \text{ and } s = t+1, \dots, T.\end{aligned}$$

Then every γ_s is universally measurable and as

$$\gamma_t + \dots + \gamma_T = q + (1 - q)(\eta_{t+1} + \dots + \eta_T) = 1$$

it follows that $\gamma \in \Gamma'_t$. Moreover,

$$d_s^\gamma = (1 - q)d_s^\eta \quad \text{for } s = t+1, \dots, T.$$

Define $Q(\omega) := P(\omega) \otimes R(\omega, \cdot)$ for $\omega \in \Omega_t$. It follows from [14, Lemma 7.29] that $Q(\omega)$ is well-defined and e.g. from [14, Lemma 7.28] and a twofold application of [14, Proposition 7.46] that $Q: \Omega_t \rightarrow \mathfrak{P}(\Omega_{T-t})$ is a kernel. Then it holds

$$\begin{aligned}(29) \quad \alpha_t(Q(\omega), \gamma(\omega, \cdot), \omega) &= \sum_{s=t}^{T-1} E_{Q(\omega)} \left[d_s^\gamma(\omega, \cdot) \beta_s \left(Q_s(\omega, \cdot), \frac{\gamma_s}{d_s^\gamma}(\omega, \cdot), (\omega, \cdot) \right) \right] \\ &= E_{P(\omega)} \left[\sum_{s=t+1}^{T-1} E_{R(\omega, \cdot)} \left[(1 - q(\omega)) d_s^\eta(\omega, \cdot) \beta_s \left(R_s(\omega, \cdot), \frac{\eta_s}{d_s^\eta}(\omega, \cdot), (\omega, \cdot) \right) \right] \right] \\ &\quad + \beta_t(P(\omega), q(\omega), \omega) \\ &= \beta_t(P(\omega), q(\omega), \omega) + (1 - q(\omega)) E_{P(\omega)} [\alpha_{t+1}(R(\omega, \cdot), \eta(\omega, \cdot), (\omega, \cdot))].\end{aligned}$$

In conclusion

$$\begin{aligned}\mathcal{E}_t(X)(\omega) &\leq q(\omega)X_t(\omega) + (1 - q(\omega))E_{P(\omega)}[\mathcal{E}_{t+1}(X1_{[t+1,T]})(\omega, \cdot)] \\ &\quad - \beta_t(P(\omega), q(\omega), \omega) + \varepsilon/2 \\ &\leq q(\omega)X_t(\omega) + (1 - q(\omega))E_{P(\omega)} \left[\sum_{s=t+1}^T E_{R(\omega, \cdot)}[\eta_s(\omega, \cdot)X_s(\omega, \cdot)] \right] \\ &\quad - \beta_t(P(\omega), q(\omega), \omega) - (1 - q(\omega))E_{P(\omega)}[\alpha_{t+1}(R(\omega, \cdot), \eta(\omega, \cdot), (\omega, \cdot))] + \varepsilon \\ &= \sum_{s=t}^T E_{Q(\omega)}[\gamma_s(\omega, \cdot)X_s(\omega, \cdot)] - \alpha_t(Q(\omega), \gamma(\omega, \cdot), \omega) + \varepsilon\end{aligned}$$

for every $\omega \in \Omega_t$ and claim (i) is established. We are left to show that (25) holds. By the above we only have to show that the left hand side in (25) is larger than the right hand side. Fix some $\omega \in \Omega_t$ and let $Q \in \mathfrak{P}(\Omega_{T-t})$ and $\gamma \in \Gamma_t$ be arbitrary. We write $Q = P \otimes R$ for a measure $P \in \mathfrak{P}(\Omega_1)$ and a kernel $R: \Omega_1 \rightarrow \mathfrak{P}(\Omega_{T-t-1})$ and define $q := \gamma_t$. Without loss of generality we may assume that $q \neq 1$, otherwise

the claim follows directly. Define $\eta_s := \gamma_s/(1-q)$ for $s = t+1, \dots, T$ so that $\eta(\bar{\omega}, \cdot) \in \Gamma_{t+1}$ for every $\bar{\omega} \in \Omega_1$. Thus, by the induction hypotheses it holds

$$\sum_{s=t+1}^T E_{R(\bar{\omega})}[\eta_s(\bar{\omega}, \cdot)X_s(\omega, \bar{\omega}, \cdot)] - \alpha_{t+1}(R(\bar{\omega}), \eta(\bar{\omega}, \cdot), (\omega, \bar{\omega})) \leq \mathcal{E}_{t+1}(X1_{[t+1, T]})(\omega, \bar{\omega})$$

for every $\bar{\omega} \in \Omega_1$. Moreover, it follows as in (29) that

$$\alpha_t(Q, \gamma, \omega) = \beta_t(P, q, \omega) + (1-q)E_{P(d\bar{\omega})}[\alpha_{t+1}(R(\bar{\omega}), \eta(\bar{\omega}, \cdot), (\omega, \bar{\omega}))]$$

from which we conclude that

$$\begin{aligned} \sum_{s=t}^T E_Q[\gamma_s(\cdot)X_s(\omega, \cdot)] - \alpha_t(Q, \gamma, \omega) &= qX_t(\omega) - \beta_t(P, q, \omega) \\ &+ (1-q)E_{P(d\bar{\omega})}\left[\sum_{s=t+1}^{T-1} E_{R(\bar{\omega})}[\eta_s(\bar{\omega}, \cdot)X_s(\omega, \bar{\omega}, \cdot)] - \alpha_{t+1}(R(\bar{\omega}), \eta(\bar{\omega}, \cdot), (\omega, \bar{\omega}))\right] \\ &\leq qX_t(\omega) + (1-q)E_{P(d\bar{\omega})}[\mathcal{E}_{t+1}(X1_{[t+1, T]})(\omega, \bar{\omega})] - \beta_t(P, q, \omega) \\ &\leq \varphi_t(X_t, \mathcal{E}_{t+1}(X1_{[t+1, T]}))(\omega) = \mathcal{E}_t(X)(\omega) \end{aligned}$$

and the proof is complete. \square

3.2. Examples. Throughout this section let $(\tilde{\varphi}_t)$ be a generator defined for random variables which is assumed to have the dual representation (9). The associated dynamic expectation $\tilde{\mathcal{E}}_t := \tilde{\varphi}_t \circ \dots \circ \tilde{\varphi}_{T-1}$ then has the representation

$$(30) \quad \tilde{\mathcal{E}}_t(X)(\omega) = \sup_{Q \in \mathfrak{P}(\Omega_{T-t})} (E_Q[X(\omega, \cdot)] - \tilde{\alpha}_t(Q, \omega)) \quad \text{for } X \in \mathcal{L},$$

where $\tilde{\alpha}_t(Q, \omega) = \sum_{s=t}^{T-1} E_Q[\tilde{\beta}_s(Q_s(\cdot), (\omega, \cdot))]$, see Proposition 2.9. The right hand side of (30) is well-defined whenever X is merely bounded and universally measurable and we shall use this equation to define $\tilde{\mathcal{E}}_t$ for such X .

3.2.1. Worst stopping. See e.g. [21, 22, 50] for the non-robust (resp. dominated) setting and e.g. [35, 47] for the continuous-time, non-dominated setting. For $t = 0, \dots, T-1$ define

$$\varphi_t(X_t, X_{t+1}) := X_t \vee \tilde{\varphi}_t(X_{t+1}) \quad \text{for } (X_t, X_{t+1}) \in \mathcal{L}_t \times \mathcal{L}_{t+1}.$$

Example 3.6. The family (φ_t) forms a generator and the associated dynamic expectation has the representation

$$(31) \quad \mathcal{E}_t(X)(\omega) = \sup_{Q \in \mathfrak{P}(\Omega_{T-t}), \tau \in \Theta_t} \left(E_Q[X_\tau(\omega, \cdot)] - E_Q\left[\sum_{s=t}^{\tau-1} \tilde{\beta}_s(Q_s(\cdot), (\omega, \cdot))\right] \right)$$

$$(32) \quad = \sup_{\tau \in \Theta_t} \tilde{\mathcal{E}}_t(X_\tau)$$

for $X \in \mathcal{R}_t$, where Θ_t is the set of all stopping times which are larger than t .

Proof. Define $\beta_s(Q, q, \omega) := \tilde{\beta}_s(Q, \omega) + \infty 1_{\{0,1\}^c}(q)$. Then (24) holds and Theorem 3.4 yields

$$\mathcal{E}_t(X)(\omega) = \sup_{Q \in \mathfrak{P}(\Omega_{T-t}), \gamma \in \Gamma_t} \left(\sum_{s=t}^T E_Q[\gamma_s(\cdot)X_s(\omega, \cdot)] - \alpha_t(Q, \gamma, \omega) \right)$$

where α_t is given by

$$\alpha_t(Q, \gamma, \omega) = \sum_{s=t}^{T-1} E_Q \left[d_s^\gamma(\cdot) \beta_s \left(Q_s(\cdot), \frac{\gamma_s(\cdot)}{d_s^\gamma(\cdot)}, (\omega, \cdot) \right) \right].$$

First notice that every stopping time τ defines an element $\gamma^\tau \in \Gamma_t$ by $\gamma_s^\tau := 1_{\tau=s}$ and it holds $d^{\gamma^\tau} = 1_{s \leq \tau-1}$. Thus the left hand side in (31) is larger than the right hand side. On the other hand, for any $\gamma \in \Gamma_t$ it must hold that $\gamma_s/d_s^\gamma \in \{0, 1\}$ Q -almost surely, that is $\gamma_s \in \{0, 1\}$ Q -almost surely, for every Q such that $\alpha_t(Q, \gamma, \omega) < +\infty$. Thus, defining the stopping time τ by $\tau := s$ on the set $\gamma_s > 0$ yields equality in (31). Further, since each β_s is positive, it follows that (31) is larger than (32). To show the other inequality, let P_s be kernels such that $\beta_s(P_s(\cdot), \cdot) \leq \varepsilon$ (see the proof of Proposition 2.9), and fix some Q and τ . Define

$$R_s := Q_s 1_{s \leq \tau-1} + P_s 1_{s \geq \tau} \quad \text{and} \quad R := R_t \otimes \cdots \otimes R_{T-1}.$$

Then $E_R[X_\tau] = E_Q[X_\tau]$ and $\tilde{\alpha}_t(R, \omega) \leq E_Q[\sum_{s=t}^{\tau-1} \tilde{\beta}_s(Q_s(\cdot), (\omega, \cdot))] + (T-t)\varepsilon$ from which we can deduce the missing inequality in (32). \square

3.2.2. Average over time. See also [21, 22]. Let $\lambda_t: \Omega_t \rightarrow (0, 1]$ be Borel such that $1/\lambda_t$ is bounded and define

$$\mathcal{E}'_t(X) := \lambda_t \tilde{\varphi}_t \circ \cdots \circ \lambda_{T-1} \tilde{\varphi}_{T-1} \left(\frac{X}{\lambda_t \cdots \lambda_{T-1}} \right) \quad \text{for } X \in \mathcal{L}$$

as well as

$$\varphi_t(X_t, X_{t+1}) := \lambda_t X_t + (1 - \lambda_t) \tilde{\varphi}_t(X_{t+1}) \quad \text{for } (X_t, X_{t+1}) \in \mathcal{L}_t \times \mathcal{L}_{t+1}.$$

Example 3.7. The family (φ_t) forms a generator and the associated dynamic expectation \mathcal{E}_t has the representation

$$\begin{aligned} \mathcal{E}_t(X)(\omega) &= \sup_{Q \in \mathfrak{P}(\Omega_{T-t})} \left(E_Q \left[\sum_{s=t}^T \gamma_s(\cdot) X_s(\omega, \cdot) \right] - \sum_{s=t}^{T-1} E_Q[d_s^\gamma(\cdot) \tilde{\beta}_s(Q_s(\cdot), (\omega, \cdot))] \right) \\ &= \mathcal{E}' \left(\sum_{s=t}^T \gamma_s X_s \right)(\omega) \end{aligned}$$

for $X \in \mathcal{R}_t$ and $\omega \in \Omega_t$, where $\gamma_t := 1 - \lambda_t$, $\gamma_s := (1 - \lambda_s) \lambda_t \cdots \lambda_{s-1}$ for $t < s < T$ and $\gamma_T := \lambda_t \cdots \lambda_{T-1}$.

Proof. Define $\beta_t(Q, q, \omega) := \tilde{\beta}_t(Q, \omega) + \infty 1_{\{\lambda_t(\omega)\}^c}(q)$ so that (24) holds. Therefore we can apply Theorem 3.4 and a computation as in Proposition 3.6 shows that γ has the claimed form. Thus the first equation holds true. Further, since γ_s is Borel and positive it holds $\sum_{s=t}^T \gamma_s X_s \in \mathcal{L}$ and since $d_t^\gamma = 1$, $d_s^\gamma = \lambda_t \cdots \lambda_{s-1}$ for $s > t$ the second equation follows. \square

3.2.3. Optimal transport for processes. Let $\Omega_1 = \mathbb{R}$ so that $\Omega = \mathbb{R}^T$ and fix a family $(\nu_t)_{1 \leq t \leq T}$ of probability measures on \mathbb{R} . For a family $(\vartheta_t)_{t \leq T}$ of bounded Borel functions from \mathbb{R} to itself and $\vartheta_0 \in \mathbb{R}$, define $I^\vartheta \in \mathcal{R}_0$ by

$$I_t^\vartheta(\omega) := \vartheta_0 + \vartheta_1(\omega_1) + \cdots + \vartheta_t(\omega_t) \quad \text{for } t = 0, \dots, T \text{ and } \omega \in \Omega_t.$$

For $X \in \mathcal{R}_0$ define

$$\mathcal{E}_0(X) := \inf \{ \vartheta_0 + E_{\nu_1}[\vartheta_1] + \cdots + E_{\nu_T}[\vartheta_T] : \vartheta \text{ such that } I^\vartheta \geq X \}.$$

Thus \mathcal{E}_0 corresponds to the minimal price needed, if one can invest statically at time $t = 0$ with outcome I_t^ϑ at time t , and the task is to super-replicate a whole given process X .

Example 3.8. *For every $X \in \mathcal{R}_0$ it holds*

$$\mathcal{E}_0(X) = \sup\{E_Q[X_T] : Q \in \mathfrak{P}(\mathbb{R}^T) : Q \circ \pi_t^{-1} = \nu_t \text{ for } t = 1, \dots, T\}$$

where $\pi_t : \mathbb{R}^T \rightarrow \mathbb{R}$ denotes the canonical projection on the t -th coordinate. Moreover, the infimum in the definition of \mathcal{E}_0 is attained.

Once measurable duality for the classical optimal transport is established (see e.g. [10, 13, 42, 48]) a simpler proof is possible – see also [52] for an extensive treatment of various aspects in optimal transport. However, our goal is to show that the present approach yields a short and structured proof, where one only has to verify the two assumptions (A) and (B).

Proof. We apply Proposition 3.5 where by Lemma 3.3 we can show (B') and (B'') instead of (B). Let $X^n \in \mathcal{R}_0^{C_b}$ be a sequence which decreases pointwise to 0. Let $\varepsilon > 0$ be arbitrary and fix some compact set $K \subset \mathbb{R}$ such that $\nu_t(K^c) \leq \varepsilon$ for every t . By Dini's lemma, there exists an index n_0 such that $X_t^n 1_{K^c} \leq \varepsilon$ for every t and $n \geq n_0$. Define $\vartheta_t := \|X^1\|_\infty 1_{K^c} + \varepsilon$. Then $X_t^n \leq \varepsilon + X_t^n 1_{(K^c)^c} \leq I_t^\vartheta$ for $n \geq n_0$ which, as $\varepsilon > 0$ was arbitrary, shows $\mathcal{E}_0(X^n) \downarrow 0$. Thus (B') follows. Let

$$a := \sup_{X \in \mathcal{R}_0^{C_b}} (E_{\bar{Q}}[X] - \mathcal{E}_0(X)) \quad \text{and} \quad b := \sup_{X \in \mathcal{R}_0^{usc_b}} (E_{\bar{Q}}[X] - \mathcal{E}_0(X))$$

for $\bar{Q} \in \mathfrak{P}(\{0, \dots, T\} \times \Omega)$. If $\bar{Q} = \delta_T \otimes Q$ for some Q such that $Q \circ \pi_t^{-1} = \nu_t$ for every t , it follows from the definition of \mathcal{E}_0 that $a = b = 0$. Otherwise define $q := \bar{Q}(\{T\} \times \Omega)$ and $Q := \bar{Q}(\{T\} \times \cdot) / q$ with any convention if $q = 0$. If $Q \circ \pi_s^{-1} \neq \nu_s$ for some s , then $E_Q[f \circ \pi_s] > E_{\nu_s}[f]$ for some $f \in C_b(\mathbb{R})$ where, without loss of generality, we may assume that f is positive. Define $X_T := x + f \circ \pi_s$ and $X_t := 0$ for $t \neq T$ as well as $\vartheta_s := f$, $\vartheta_T := x$ and $\vartheta_t := 0$ for $t \notin \{s, T\}$, where $x \in \mathbb{R}$ is arbitrary. Then $I^\vartheta \geq X$ so that

$$E_{\bar{Q}}[X] - \mathcal{E}_0(X) \geq qE_Q[X_T] - (E_{\nu_s}[f] + x) = x(q - 1) + qE_Q[f \circ \pi_s] - E_{\nu_s}[f].$$

Scaling in x yields $a = b = +\infty$ if $q \neq 1$ and then scaling in f shows $a = b = +\infty$ if $Q \circ \pi_s^{-1} \neq \nu_s$. Therefore (B'') is established.

We are left to show (A). Let $X^n \in \mathcal{R}_0$ such that $X^n \uparrow X \in \mathcal{R}_0$ where, since $\mathcal{E}_0(Y + 1_{[0, T]c}) - c = \mathcal{E}_0(Y)$ for $Y \in \mathcal{R}_0$, we may assume that $X^n \geq 0$. Further let ϑ^n be an optimal strategy for $\mathcal{E}_0(X^n)$ within an error of $1/n$. Then, since

$$0 \leq \inf_{\omega \in \Omega} X_T(\omega) \leq \inf_{\omega \in \Omega} I_T^{\vartheta^n}(\omega) = \sum_t \inf_{x \in \mathbb{R}} \vartheta_t^n(x),$$

there exist $\eta_t^n \leq \inf_x \vartheta_t^n(x)$ such that $\sum_t \eta_t^n = 0$. Now define $\zeta_t^n := \vartheta_t^n - \eta_t^n + \|X\|_\infty$ for $t = 0, \dots, T-1$ and $\zeta_T^n := \vartheta_T^n - \eta_T^n - T\|X\|_\infty$. Then $I^{\zeta^n} \geq X^n$ and $\sum_t E_{\nu_t}[\zeta_t^n] = \sum_t E_{\nu_t}[\vartheta_t^n]$. Moreover, every ζ_t^n is bounded from below by $-T\|X\|_\infty$ and $E_{\nu_t}[\zeta_t^n]$ is bounded since $\mathcal{E}_0(X^n)$ is bounded by $\|X\|_\infty$. Therefore, after passing to forward convex combinations, we may assume that ζ_t^n has a ν_t -almost sure limit ζ_t , see [30, Lemma A1.1]. Define $\vartheta_t := \zeta_t$ on the set where ζ_t^n converges, and $+\infty$ else. Then $I^\vartheta \geq X$ and Fatou's lemma yields $\mathcal{E}_0(X) \leq \lim_n \mathcal{E}_0(X^n)$ which, by monotonicity of \mathcal{E}_0 , shows (A). For the existence of an optimal strategy ϑ , one applies the previous argumentation to the constant sequence $X_n := X$. \square

APPENDIX A. ANALYTIC SETS

In the sequel let V and W be two Polish spaces. Recall that a subset of a Polish space is called analytic if it is the image of a Borel set of another Polish space under a Borel function. A function $f: V \rightarrow \mathbb{R} \cup \{\pm\infty\}$ is upper semianalytic, if $\{f \geq c\} \subset V$ is an analytic set for every real number c and lower semianalytic, if $-f$ is upper semianalytic. The set of universally measurable subsets of V is defined as $\bigcap \{\mathcal{B}(V)^P : P \in \mathfrak{P}(V)\}$, where $\mathcal{B}(V)^P$ is the completion of the Borel σ -field $\mathcal{B}(V)$ with respect to the probability P . A function $f: V \rightarrow W$ is said to be universally measurable, if $\{f \in B\}$ is universally measurable for every $B \in \mathcal{B}(W)$. It follows from the definition that every Borel set is analytic, and from Lusin's theorem (see [14, Proposition 7.42]) that every analytic set is universally measurable. The same of course holds true if we replace sets by functions in the previous sentence. Finally, we say that a set-valued mapping $\Phi: V \rightsquigarrow W$ has analytic graph, if

$$\text{graph } \Psi := \{(v, w) : w \in \Psi(v)\} \subset V \times W$$

is an analytic set. For a countable family $\{X_n : n\}$ of upper semianalytic functions, $X_1 + X_2$, $\sup_n X_n$, and $\inf_n X_n$ are again upper semianalytic. Moreover, if Y is a positive Borel function, then YX_1 is upper semianalytic (see e.g. [14, Lemma 7.30] for both statements). A comprehensive treatment of analytic sets can be found e.g. in [14, Chapter 7].

APPENDIX B. REPRESENTATION RESULTS

The following theorem is a slight generalization of [8, Theorem A.1] in that we consider a more general space of functions instead of the continuous bounded ones, however, the proof can be copied. Let Ω be Polish space. For a set M of functions $X: \Omega \rightarrow \mathbb{R}$ we denote by M_δ all functions which can be written as the limit of a decreasing sequence of functions in M . Further the set of M -Soulin functions is denoted by $S(M)$, that is, $S(M)$ consists of all functions X which can be written as $X = \sup_{\gamma \in \mathbb{N}^\mathbb{N}} \inf_{n \in \mathbb{N}} \sigma(\gamma_1, \dots, \gamma_n)$ for a mapping $\sigma: \bigcup \{\mathbb{N}^n : n \in \mathbb{N}\} \rightarrow M$. Finally $M_{\delta,b}$ and $S(M)_b$ denotes the set of all bounded functions in M_δ respectively $S(M)$ and $\mathfrak{M}(M)$ the set of all positive measures on the σ -field generated by M .

Theorem B.1. *Let M be a linear space of bounded functions from Ω to \mathbb{R} such that $1 \in M$ and $X \vee Y \in M$ whenever $X, Y \in M$. Let $\phi: S(M) \rightarrow \mathbb{R}$ be an increasing convex functional which satisfies*

- (1) $\phi(X_n) \downarrow 0$ for every sequence $X_n \in M$ such that $X_n \downarrow 0$,
 - (2) $\phi^*(Q) := \sup_{X \in M} (E_Q[X] - \phi(X)) = \sup_{X \in M_{\delta,b}} (E_Q[X] - \phi(X))$ for all $Q \in \mathfrak{M}(M)$,
 - (3) $\phi(X_n) \uparrow \phi(X)$ for every sequence $X_n \in S(M)_b$ such that $X_n \uparrow X \in S(M)_b$.
- Then $\Lambda_c := \{Q \in \mathfrak{M}(M) : \phi^*(Q) \leq c\}$ is $\sigma(\mathfrak{M}(M), M)$ compact and

$$\phi(X) = \sup_{Q \in \mathfrak{M}(M)} (E_Q[X] - \phi^*(Q))$$

for all $X \in S(M)_b$.

Lemma B.2. *Let Ω_t be as in Section 2 and let $\bar{\Omega} := \{t, \dots, T\} \times \Omega_T$. Let $M := C_b(\Omega_t)$ resp.*

$$M = \{X \in C_b(\bar{\Omega}) : X_s(\omega) \text{ only depends on } (\omega_1, \dots, \omega_s) \text{ for } s = t, \dots, T\}$$

Then $S(M)_b = \text{usa}_b(\Omega_t)$ resp.

$$S(M)_b = \{X: \bar{\Omega} \rightarrow \mathbb{R} : X_s \in \text{usa}_b(\Omega_s) \text{ and } X_s(\omega) \text{ only depends on } (\omega_0, \dots, \omega_t)\}.$$

Proof. The first claim is proven e.g. in [8, Theorem A.1]. The second claim follows from the definition of Souslin functions and the observation that $X: \bar{\Omega} \rightarrow \mathbb{R}$ is upper semianalytic if and only if X_s is upper semianalytic for $s = t, \dots, T$. \square

REFERENCES

- [1] B. Acciaio, M. Beiglböck, F. Penkner, and W. Schachermayer. A model-free version of the fundamental theorem of asset pricing and the super-replication theorem. *Mathematical Finance*, 2013.
- [2] B. Acciaio, H. Föllmer, and I. Penner. Risk assessment for uncertain cash flows: Model ambiguity, discounting ambiguity, and the role of bubbles. *Finance and Stochastics*, 16(4):669–709, 2012.
- [3] B. Acciaio and I. Penner. Dynamic risk measures. In *Advanced Mathematical Methods for Finance*, pages 1–34. Springer, 2011.
- [4] C. Aliprantis and K. Border. *Infinite dimensional analysis: a hitchhiker’s guide*. Springer Science & Business Media, 2006.
- [5] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.
- [6] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath, and H. Ku. Coherent multiperiod risk adjusted values and bellmans principle. *Annals of Operations Research*, 152(1):5–22, 2007.
- [7] D. Bartl. Exponential utility maximization under model uncertainty for unbounded endowments. *arXiv preprint arXiv:1610.00999*, 2016.
- [8] D. Bartl, P. Cheridito, and M. Kupper. Robust expected utility maximization with medial limits. *forthcoming*.
- [9] E. Bayraktar and Y. Zhang. Fundamental theorem of asset pricing under transaction costs and model uncertainty. *Mathematics of Operations Research*, 41(3):1039–1054, 2016.
- [10] M. Beiglböck, M. Goldstern, G. Maresch, and W. Schachermayer. Optimal and better transport plans. *Journal of Functional Analysis*, 256(6):1907–1927, 2009.
- [11] M. Beiglböck, P. Henry-Labordère, and F. Penkner. Model-independent bounds for option prices – a mass transport approach. *Finance and Stochastics*, 17(3):477–501, 2013.
- [12] M. Beiglböck, M. Nutz, and N. Touzi. Complete duality for martingale optimal transport on the line. *The Annals of Probability forthcoming*, *arXiv:1507.00671*, 2015, 2015.
- [13] M. Beiglböck and W. Schachermayer. Duality for borel measurable cost functions. *Transactions of the American Mathematical Society*, 363(8):4203–4224, 2011.
- [14] D. Bertsekas and S. Shreve. *Stochastic optimal control: The discrete time case*, volume 23. Academic Press New York, 1978.
- [15] B. Bouchard and M. Nutz. Arbitrage and duality in nondominated discrete-time models. *The Annals of Applied Probability*, 25(2):823–859, 2015.
- [16] N. Bourbaki and S. K. Berberian. *Integration II*. Springer, 2004.
- [17] M. Burzoni, M. Frittelli, and M. Maggis. Model-free superhedging duality. *The Annals of Applied Probability forthcoming*, *arXiv:1506.06608*, 2015.
- [18] L. Carassus and R. Blanchard. Robust optimal investment in discrete time for unbounded utility function. *arXiv preprint arXiv:1609.09205*, 2016.
- [19] P. Cheridito, F. Delbaen, and M. Kupper. Coherent and convex monetary risk measures for bounded cadlag processes. *Stochastic Processes and their Applications*, 112(1):1–22, 2004.
- [20] P. Cheridito, F. Delbaen, and M. Kupper. Coherent and convex monetary risk measures for unbounded cadlag processes. *Finance and Stochastics*, 9(3):369–387, 2005.
- [21] P. Cheridito, F. Delbaen, and M. Kupper. Dynamic monetary risk measures for bounded discrete-time processes. *Electronic Journal of Probability*, 11(3):57–106, 2006.
- [22] P. Cheridito and M. Kupper. Composition of time-consistent dynamic monetary risk measures in discrete time. *International Journal of Theoretical and Applied Finance*, 14(01):137–162, 2011.
- [23] P. Cheridito, M. Kupper, and L. Tangpi. Representation of increasing convex functionals with countably additive measures. *arXiv preprint arXiv:1502.05763*, 2015.

- [24] P. Cheridito, M. Kupper, and L. Tangpi. Duality formulas for robust pricing and hedging in discrete time. *arXiv preprint arXiv:1602.06177*, 2016.
- [25] G. Choquet. Theory of capacities. In *Annales de l'institut Fourier*, volume 5, pages 131–295. Institut Fourier, 1954.
- [26] G. Choquet. Forme abstraite du théorème de capacitabilité. In *Annales de l'institut Fourier*, volume 9, pages 83–89, 1959.
- [27] S. Cohen. Quasi-sure analysis, aggregation and dual representations of sublinear expectations in general spaces. *Electron. J. Probab*, 17(62):1–15, 2012.
- [28] F. Delbaen. Coherent risk measures on general probability spaces. In *Advances in finance and stochastics*, pages 1–37. Springer, 2002.
- [29] F. Delbaen. The structure of m -stable sets and in particular of the set of risk neutral measures. In *In Memoriam Paul-André Meyer*, pages 215–258. Springer, 2006.
- [30] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische annalen*, 300(1):463–520, 1994.
- [31] C. Dellacherie and P.-A. Meyer. *Probabilities and Potential, C: Potential Theory for Discrete and Continuous Semigroups*, volume 151. Elsevier, 2011.
- [32] L. Denis and M. Kervarec. Optimal investment under model uncertainty in nondominated models. *SIAM Journal on Control and Optimization*, 51(3):1803–1822, 2013.
- [33] K. Detlefsen and G. Scandolo. Conditional and dynamic convex risk measures. *Finance and Stochastics*, 9(4):539–561, 2005.
- [34] I. Ekren and H. M. Soner. Constrained optimal transport. *arXiv preprint arXiv:1610.02940*, 2016.
- [35] I. Ekren, N. Touzi, and J. Zhang. Optimal stopping under nonlinear expectation. *Stochastic Processes and Their Applications*, 124(10):3277–3311, 2014.
- [36] K. Fan. Minimax theorems. *Proceedings of the National Academy of Sciences*, 39(1):42–47, 1953.
- [37] H. Föllmer and A. Schied. *Stochastic finance: an introduction in discrete time*. Walter de Gruyter, 2011.
- [38] M. Frittelli and E. R. Gianin. Dynamic convex risk measures. *Risk measures for the 21st century*, pages 227–248, 2004.
- [39] M. Frittelli and G. Scandolo. Risk measures and capital requirements for processes. *Mathematical finance*, 16(4):589–612, 2006.
- [40] D. G. Hobson. Robust hedging of the lookback option. *Finance and Stochastics*, 2(4):329–347, 1998.
- [41] M. Katětov. On real-valued functions in topological spaces. *Fundamenta Mathematicae*, 38(1):85–91, 1951.
- [42] H. Kellerer. Duality theorems for marginal problems. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 67(4):399–432, 1984.
- [43] D. Lacker. A non-exponential extension of sanov’s theorem via convex duality. *arXiv preprint arXiv:1609.04744*, 2016.
- [44] M. Maggis, T. Meyer-Brandis, and G. Svindland. The fatou property under model uncertainty. *arXiv preprint arXiv:1610.04085*, 2016.
- [45] A. Neufeld and M. Šikić. Robust utility maximization in discrete-time markets with friction. *arXiv preprint arXiv:1610.09230*, 2016.
- [46] M. Nutz and R. van Handel. Constructing sublinear expectations on path space. *Stochastic Processes and their Applications*, 123(8):3100–3121, 2013.
- [47] M. Nutz and J. Zhang. Optimal stopping under adverse nonlinear expectation and related games. *The Annals of Applied Probability*, 25(5):2503–2534, 2015.
- [48] S. T. Rachev and L. Rüschendorf. *Mass Transportation Problems: Volume I: Theory*, volume 1. Springer Science & Business Media, 1998.
- [49] F. Riedel. Dynamic coherent risk measures. *Stochastic processes and their applications*, 112(2):185–200, 2004.
- [50] F. Riedel. Optimal stopping with multiple priors. *Econometrica*, 77(3):857–908, 2009.
- [51] D. Stroock. *Probability theory: an analytic view*. Cambridge university press, 2010.
- [52] C. Villani. *Optimal transport: old and new*, volume 338. Springer Science & Business Media, 2008.